Annals of Pure and Applied Logic 64 (1993) 95–152
North-Holland

Subsystems of true arithmetic and hierarchies of functions

Z. Ratajczyk
Instytut Matematyki, Universytet Warszawski, Banacha 2, 02-097 Warszawa, Poland

Communicated by D. van Dalen
Received 14 November 1989
Revised 11 January 1993

Abstract

Ratajczyk, Z., Subsystems of true arithmetic and hierarchies of functions, Annals of Pure and Applied Logic 64 (1993) 95–152.

The combinatorial method coming from the study of combinatorial sentences independent of PA is developed. Basing on this method we present the detailed analysis of provably recursive functions associated with higher levels of transfinite induction, $I(\varepsilon_\alpha)$, and analyze combinatorial sentences independent of $I(\varepsilon_\alpha)$. Our treatment of combinatorial sentences differs from the one given by McAloon [18] and gives more natural sentences. The same method give also a combinatorial technique with no use of the cut-elimination theorem which is appropriate to study proof-theoretic strength of subsystems of first order arithmetic and some of their expansions. It was used to analyze iterated reflection principle and system of transfinite induction with a satisfaction class.

Contents

Correspondence to: Z. Ratajczyk, Instytut Matematyki, Universytet Warszawski, Banacha 2, 02-097 Warszawa, Poland.

0. Introduction

The aim of this paper is to give some applications of combinatorial methods in the proof theory of subsystems of true arithmetic such as arithmetical transfinite induction $\mathrm{I}(\varepsilon_\alpha)$, and the iterated reflection principle $\mathcal{R}_\alpha(\mathrm{PA})$. These methods come out from the study of combinatorial-independent (from PA) statements of the Ramsey style.

For explanation, $\mathcal{R}_0(\mathrm{PA}) = \mathrm{PA} + \mathrm{RFN}(\mathrm{PA})$. One of the definitions of the iterated reflection principles appears in Feferman [4] (see also Smorynski [30]). We restate this definition in Chapter IV. The principle $\mathrm{I}(\varepsilon_\alpha)$, denoted in the literature also as $\mathrm{TI}(\varepsilon_\alpha)$, is the theory of transfinite induction up to ε_α; this is well defined if we are given a primitive recursive system of notations for some ordinal $\lambda \geq \varepsilon_\alpha$ and the operations $\beta + \gamma$, ω^β and ε_β are defined in PA.

The aim of this paper is twofold. On one hand we construct some simple combinatorial sentences in the Ramsey style independent from the systems mentioned above. On the other hand we show that the same methods which allow us to construct these statements give also a proof-theoretic technique yielding also other proof-theoretic results with no use of the cut-elimination theorem.

To be more precise, we generalize the Paris–Harrington statement [21] to the $\mathrm{I}(\varepsilon_\alpha)$-case. Moreover, we generalize to the $\mathrm{I}(\varepsilon_\alpha)$-case Wainer's theorem which describes the set Rek(PA)—the recursive functions provably total in PA, in terms of the Hardy hierarchy up to ε_0. We also prove the equivalence $\mathrm{I}(\varepsilon_\alpha) \equiv \mathcal{R}_\alpha(\mathrm{PA})$.

The starting point for the generalizations of Wainer's result is the combinatorial proof of this theorem for PA, see [20] and [9], see also [23], where a short proof is given. Some of the technical details come from [23].

We stress that the classical proof-theoretic argument for Wainer's theorem is quite long. Wainer's paper [32] is just the final step of the proof based on the description of Rek(PA) in terms of $<\varepsilon_0$ recursion (known previously, see [13] and [31]). A much shorter and elegant proof-theoretic argument was found quite recently, see [2].

Let us also add that the main tool for the description of $\mathrm{Rek}(\mathrm{I}(\varepsilon_\alpha))$ used below is the Hardy hierarchy of any primitive recursive length λ. The definition of this hierarchy depends on the choice of the family of sequences $\alpha_n \nearrow \alpha$, for each limit $\alpha < \lambda$. We denote this family by P and call it the system of (fundamental) sequences. The Hardy's hierarchy constructed with the use of P is denoted by H_α^P, $\alpha < \lambda$. We shall need a very special system P of sequences; see the discussion of this topic in Chapter I.

We have organized the paper as follows. Chapter I is devoted to some aspects of the fundamental sequences and Hardy's hierarchies. Chapter II contains the main combinatorial and logical notions used in the paper. The other chapters are devoted to the proofs of the main results of the paper.

Let us list the informal statements of the main results.

In Chapter III we obtain some results connected with Wainer's theorem. The main ones are:

Corollary to Theorem III.1.1 (immediate generalization of Wainer's theorem). *If P is an appropriate system of sequences for $\varepsilon_{\alpha+1}$ then the following implication holds:*

$$f \in \mathrm{Rek}(\mathrm{I}(\varepsilon_\alpha)) \to \exists \beta < \varepsilon_{\alpha+1}\, \exists x\, \forall y > x\, f(x) < H^P_\beta(y).$$

Let P-$\mathrm{TH}^0(<\alpha)$ denote $\mathrm{Cn}(\mathrm{I}\Sigma_1) \cap \Pi_2$ plus all the axioms

$$\forall x\, \exists y\, H^P_\beta(x) = y \quad \text{for } \beta < \alpha.$$

Theorem III.1.3. *If P is an appropriate system of sequences for $\varepsilon_{\alpha+1}$, then $(\mathrm{I}(\varepsilon_\alpha) \restriction \Pi_2) \equiv P$-$\mathrm{TH}^0(<\varepsilon_{\alpha+1})$, where $\mathrm{I}(\varepsilon_\alpha) \restriction \Pi_2$ denotes the theory $\mathrm{Cn}(\mathrm{Cn}(\mathrm{I}(\varepsilon_\alpha)) \cap \Pi_2)$.*

In Chapter IV we present a new proof of the Kreisel–Levy–Schmerl theorem [14, 27]), $\mathscr{R}_\alpha(\mathrm{PA}) \equiv \mathrm{I}(\varepsilon_\alpha)$.

In order to state the next results of Chapter IV, let PH^0 denote the principle PH which Paris and Harrington [21] proved to be independent of PA. Moreover let PH^α_Q denote the appropriate αth iteration of PH, depending on the system Q of sequences. The principle PH^α_Q defined in this paper is a refinement of the combinatorial principle studied by McAloon [18].

Theorem IV.2.1. *If Q is the appropriate system of sequences for λ in $\mathrm{I}\Sigma_1$, Q is Σ_1 then $\mathrm{I}\Sigma_1 \vdash \mathrm{PH}^\alpha_Q \equiv \mathrm{R}(\mathrm{I}(<\varepsilon_\alpha), \Sigma_1)$, where $\mathrm{R}(T; \Sigma_1)$ denotes the uniform reflection principle for Σ_1 formulas with respect to T.*

In Chapter V we go beyond the first-order arithmetic. Let $\mathrm{I}(\varepsilon_\alpha; S)$ denote PA plus transfinite induction up to ε_α in the language $L_{\mathrm{PA}}(S)$ plus the sentence stating that S is a full satisfaction class for L_{PA}-formulas.

Theorem V.1. *If P is an appropriate system of sequences for $\varepsilon_{\varepsilon_{\alpha+1}}$ then $\mathrm{I}(\varepsilon_\alpha; S) \restriction \Pi_2 \equiv P$-$\mathrm{TH}^0(<\varepsilon_{\varepsilon_{\alpha+1}})$.*

We conclude this introduction with a description of the main notions and ideas used in this paper. We extensively work with sets of diagonally indiscernible sets in the Paris–Harrington sense [21]. The central role is played by the notion of some relatively large sets of diagonal indiscernibles, called a-skeletons (see Definition II.1.4). An important idea is to use them in a specific manner. Namely for an a-skeleton A we define the relation $A \vDash \theta$ for $\theta < a$, $\theta \in L_{\mathrm{PA}}$, which has some properties similar to the usual satisfaction relation (Lemma II.1.9). This relation is adequate for the limited provability (Lemma II.1.10). Moreover,

$$\forall \theta \in \mathrm{Ax}(\mathrm{PA}) \cap [0, a]\;\; A \vDash \theta.$$

The strength of the theories $P\text{-TH}^0(<\varepsilon_{\alpha+1})$, $\mathrm{I}(\varepsilon_{\alpha+1})$ and $\mathrm{PA}+\mathrm{PH}^{\alpha}_{Q}$ allows us to construct some very large a-skeletons. Then we show that there exist a-skeletons which are a-models for $\mathrm{I}(\varepsilon_\alpha)$. Lemma III.1.5, the main lemma used in the proof of Theorem III.1.1 is roughly speaking as follows.

$\mathrm{I}\Sigma_1$ proves: if c is sufficiently large then every ε_α-large (in the Ketonen–Soloway sense) a-skeleton A has the following property:

$$\forall\theta\,[\mathrm{Ind}(\theta,\,\varepsilon_\alpha)<a^{1/c}\rightarrow A\vDash\mathrm{Ind}(\theta,\,\varepsilon_\alpha)].$$

The construction of a-skeletons is more or less in the style of the proof of the Paris–Harrington result. They are not used to construct inner models, however. Indeed, the transfinite iteration of the construction of such models does not seem possible. Rather than this we study immediate connections similar to the one expressed in the lemma stated above. In the proof of the lemma we formally argue as we would have the transfinite induction on α. This is admissible because such induction is in fact the finite induction thanks to the possibility of the so-called finitization of this sort of reasoning (see Lemma I.2.6).

The method of the proof described above is called combinatorical-logical (the simplest case of this method appears in Chapter II). This method was used in the proofs of all the main results. Its advantage over the classical cut-elimination technique is that it gives simple information about the change of the lengths of the proofs if we pass from a fragment of one theory to the equivalent fragment of another theory.

Acknowledgements

I would like to thank all my colleagues for valuable comments on the paper, especially Z. Adamowicz who read the manuscript. Thanks are also due to J. Trzeciak for correcting the language. I am most grateful to my wife, Barbara, for her patience and help.

I. Systems of fundamental sequences and hierarchies

In this chapter we consider the problem of choice of a system P of fundamental sequences $\alpha_n\nearrow\alpha$ for each limit $\alpha<\lambda$ (for λ for short), where λ is a fixed primitive recursive ordinal.

We make precise conditions which ensure among other things that all functions in the hierarchy H^P_α: $\alpha<\lambda$ are increasing. The formal arithmetical definition of the hierarchy is in 2.5. Set-theoretically the hierarchy H^P_α: $\alpha<\lambda$ is defined by inductive conditions:

$$H^P_0=\mathrm{id},$$
$$H^P_{\alpha+1}(x)\simeq H^P_\alpha(x+1)\quad\text{for } x\in\mathbb{N},$$
$$H^P_\alpha(x)\simeq H^P_{\alpha_x}(x)\quad\text{for } \alpha\in Lim,\ x\in\mathbb{N}.$$

Here the symbol $\simeq$ denotes the alternative: are identical or both undefined. In order to ensure that the hierarchy H_α^P: $\alpha<\lambda$ consists only of increasing functions the following condition (cf. [17]) is sufficient:

(*) $\qquad H_{\alpha_i}^P(x) < H_{\alpha_x}^P(x)$ for all $i<x\in\mathbb{N}$ and all $\alpha<\lambda$, $\alpha\in Lim$.

Indeed, if assume (*) then the inductive step for $\alpha\in Lim$ follows from the following inequalities for $i<x$:

$$H_\alpha^P(i) = H_{\alpha_i}^P(i) < H_{\alpha_i}^P(x) < H_{\alpha_x}^P(x) = H_\alpha^P(x).$$

The condition expressed by (*) will be called local monotonicity of the hierarchy.

Schmidt [28, 29] proved that if the system P of sequences satisfies the following condition due to Bachmman [1]

(**) $\qquad$ for $\alpha,\beta<\lambda$ and $n\in\mathbb{N}$ we have $\quad \alpha_n<\beta<\alpha \rightarrow \alpha_n\leqslant\beta_0$,

then each well-defined hierarchy (H_α^P: $\alpha<\lambda$ is such an example) is locally monotone, and hence contains only increasing functions. Moreover the hierarchy itself is increasing, i.e., $\alpha<\beta$ implies $\exists y\,\forall x>y\, H_\alpha^P(x)<H_\beta^P(x)$. Systems of (fundamental) sequences satisfying (**) are called B-systems. A simple example of a B-system is obtained from the usual one for ε_0 by a slight modification $[(\omega^{\alpha+1})_n=\omega^\alpha(n+1)$ rather than $\omega^\alpha n]$, cf. [12, 2.4].

In Section 1 we present formal arithmetic (indeed, in $\mathrm{I}\Sigma_1$) counterparts of the following notions: system of notations, system of ε-notation, system of sequences and B-system of sequences. We define also a bit stronger notion of a B^+-system and also of a B^+-ε-system of sequences.

In Section 2 we prove that most of the above-mentioned properties of the hierarchy H_α^P: $\alpha<\lambda$ are provable in $\mathrm{I}\Sigma_1$, provided P is a B^+-system in $\mathrm{I}\Sigma_1$. Moreover we prove in Section 2 the existence of B^+-systems and of B^+-ε-systems in $\mathrm{I}\Sigma_1$, for every primitive recursive λ.

1.1. The arithmetical definitions of systems, and hierarchies

We formalize in the theory of mathematical induction for Σ_1 formulas, $\mathrm{I}\Sigma_1$, all the basic notions concerning systems. As we shall see later all the above-mentioned notions are equivalent to analagous notions in PRA, indeed, the theorem of Minc/Takeuti says that $\mathrm{PRA}\cap\Pi_2=\mathrm{I}\Sigma_1\cap\Pi_2$. We have chosen $\mathrm{I}\Sigma_1$ because proofs in $\mathrm{I}\Sigma_1$ are less troublesome.

1.1. Definition. We say that the formula $\lambda(x,y)\in\Sigma_1$ defines the *basic system of notations in* $\mathrm{I}\Sigma_1$ iff $\lambda(x,y)$ defines in $\mathrm{I}\Sigma_1$ an unbounded discrete strict linear order on the domain of λ with least element 0, and the formulas

$$\lambda(x,y)\wedge\neg\exists z\,(\lambda(x,z)\wedge\lambda(z,y)),$$
$$\forall y\,[\lambda(y,x)\rightarrow\exists z\,(\lambda(y,z)\wedge\lambda(z,x))]$$

are of class Σ_1 in $\mathrm{I}\Sigma_1$.

The first of these formulas we shall denote as $y = S(x)$ and the second as $Lim(x)$. Obviously these formulas are automatically Π_1 in $\mathrm{I}\Sigma_1$.

We use the following conventions which will be used throughout the paper. Instead of x belongs to the domain, of λ, i.e., instead of $\exists y\,(\lambda(x, y) \vee \lambda(y, x))$, we write $x < \lambda$. The variables whose range is limited to $\operatorname{dom}\lambda$ are denoted suggestively by $\alpha, \beta, \ldots$. Instead of $\lambda(\alpha, \beta)$ we write $\alpha < \beta$. If we wish to write that α is smaller than β in the sense of the ordering on the natural numbers then we write $\ulcorner\alpha\urcorner < \ulcorner\beta\urcorner$. Obviously there exists a Σ_1 formula which defines iterations $S^n(\alpha)$ in $\mathrm{I}\Sigma_1$. We write $\beta = \alpha + n$ rather than $\beta = S^n(\alpha)$; the sum of α and n as natural numbers will be denoted as $\ulcorner\alpha\urcorner + n$. The operation of addition of ordinals will be denoted $\alpha + \beta$, whenever defined.

It is known that for every recursive ordinal α there exists a formula $\lambda(x, y) \in \Sigma_1$ satisfying $(\alpha, <) \simeq (\mathbb{N}, \{(m, n) : \mathbb{N} \models \lambda(m, n)\})$ which defines the basic system of notation in $\mathrm{I}\Sigma_1$.

Let the notation $\varphi(x, y) = z \in \Sigma_1$ mean that φ is a formula having 3 free variables ($\varphi := \varphi(x, y, z) \in \Sigma_1$) such that the sentence $\varphi(x, y, z_1) \wedge \varphi(x, y, z_2) \rightarrow z_1 = z_2$ is provable in $\mathrm{I}\Sigma_1$.

1.2. Definition. We say that the formula $P(x, y) = z \in \Sigma_1$ defines a *system of fundamental sequences for* λ *in* $\mathrm{I}\Sigma_1$ (shortly the system of sequences for λ in Σ_1) iff the theory $\mathrm{I}\Sigma_1$ proves

(i) $P(0, n) = 0 \wedge \forall\alpha > 0\ P(\alpha, n) < \alpha$,

(ii) $P(\alpha + 1, n) = \alpha$,

(iii) $Lim(\alpha) \wedge \beta < \alpha \rightarrow \exists n\ \beta < P(\alpha, n)$.

(iv) $n < m \rightarrow P(\alpha, n) < P(\alpha, m)$.

The value $P(\alpha, n)$ will also be denoted by α_n^P.

1.3. Definition. 1. We say that the formula $P(x, y) = z \in \mathrm{I}\Sigma_1$ defines a *B-system of sequences for* λ *in* $\mathrm{I}\Sigma_1$ iff

(i) P defines the system of fundamental sequences for λ in $\mathrm{I}\Sigma_1$,

(ii) $\mathrm{I}\Sigma_1 \vdash \alpha_n^P < \beta < \alpha \rightarrow \alpha_n^P \leqslant \beta_0^P$.

2. We say that P defines a *B^+-system of sequences for* λ *in* $\mathrm{I}\Sigma_1$ iff P satisfies (i) but instead of (ii), $\mathrm{I}\Sigma_1$ proves

$$\alpha_n^P + 1 < \beta < \alpha \;\rightarrow\; \alpha_n^P + 1 \leqslant \beta_0^P.$$

A simple example of a B^+-system is the system for ε_0 defined by the following conditions

$$(0)n = 0,$$

$$(\omega^{\alpha+1})_n = \omega^\alpha(n + 1) + 1,$$

$$(\omega^\alpha)_n = \omega^{\alpha_n} \quad \text{for } \alpha \in Lim.$$

It is easy to show that this system is Σ_1-definable in $\mathrm{I}\Sigma_1$.

Systems of notations λ for which the operations ω^{α}, ε_{α}, $\alpha \dotplus \beta$ are defined in primitive recursive manner are called ε-systems of notations. Here $\alpha \dotplus \beta$ is the restriction of addition to arguments α, β such that all the exponents in the Cantor normal expansion of α are $\geqslant$ all the exponents of β.

In order to work with such systems we must assume that some axiomatically defined properties of the above-mentioned operations are provable in $\mathrm{I}\Sigma_1$. In particular the axioms should ensure the Cantor normal form and continuity of operations ω^{α}, ε_{α}. We need very special ε-systems (and fundamental sequences), which guarantees nice properties of α-large sets, see Section II.3. These requirements lead to the following definition.

1.4. Definition. We say that $\lambda(x, y) \in \Sigma_1$ together with the formulas $\alpha \dotplus \beta = \gamma$, $\omega^{\alpha} = \beta$, $\varepsilon_{\alpha} = \beta \in \Sigma_1$ defines an *ε-system of notation in* $\mathrm{I}\Sigma_1$ iff $\lambda(x, y)$ defines a basis system of notation in $\mathrm{I}\Sigma_1$ and $\mathrm{I}\Sigma_1$ proves:

(i) $\alpha \dotplus 0 = 0 \dotplus \alpha = \alpha$, $\alpha \dotplus 1 = \alpha + 1$, $\omega^0 = 1$, $\omega^{\varepsilon_\alpha} = \varepsilon_{\alpha}$,

(ii) $\alpha \dotplus (\beta \dotplus \gamma) = (\alpha \dotplus \beta) \dotplus \gamma$,

(iii) $\alpha \dotplus \omega^{\beta}\downarrow \rightarrow [(\alpha \dotplus \omega^{\beta}) \dotplus \omega^{\gamma}\downarrow \Leftrightarrow \beta \leqslant \gamma]$, where the symbol $\downarrow$ indicates that the operation is defined,

(iv) $\forall \alpha \neq 0\, \exists! \beta, \gamma\, (\alpha = \beta \dotplus \omega^{\gamma})$ and $\sim \exists \beta, \gamma\, (0 = \beta \dotplus \omega^{\gamma})$,

(v) $\alpha \dotplus \beta\downarrow \wedge\ \alpha, \beta \neq 0 \rightarrow \ulcorner \alpha \urcorner, \ulcorner \beta \urcorner < \ulcorner \alpha \dotplus \beta \urcorner$,

(vi) $\neg \exists \gamma\, (\beta = \varepsilon_{\gamma}) \rightarrow \ulcorner \beta \urcorner < \ulcorner \omega^{\beta} \urcorner$,

(vii) $\alpha \dotplus \beta\downarrow \rightarrow [(\gamma < \alpha \vee \exists \delta < \beta\ \gamma = \alpha \dotplus \delta) \rightarrow \gamma < \alpha \dotplus \beta]$,

(viii) $\omega^{\gamma}\downarrow \rightarrow (\omega^{\alpha} \dotplus \beta < \omega^{\gamma} \Leftrightarrow \alpha < \gamma)$,

(ix) $\varepsilon_{\gamma}\downarrow \rightarrow (\varepsilon_{\beta} < \varepsilon_{\gamma} \Leftrightarrow \beta < \gamma)$.

ε-systems play an essential role in the studies of arithmetical transfinite induction $\mathrm{I}(\alpha)$. For completeness let us formulate an exact formal definition of $\mathrm{I}(\alpha)$ mentioned in the Introduction. Assume that λ defines a basic system of notations in $\mathrm{I}\Sigma_1$. Let $\underset{\sim}{\alpha}$ be a term defining number α such that $\mathrm{I}\Sigma_1 \vdash \underset{\sim}{\alpha} < \lambda$.

1.5. Definition. The theory $\mathrm{I}(\alpha)$ is PA extended by the following scheme of arithmetic transfinite induction:

$$\forall y\ \{\forall \beta < \underset{\sim}{\alpha}\ [\forall \gamma < \beta\ \varphi(\gamma, y) \rightarrow \varphi(\beta, y)] \rightarrow \forall \beta < \underset{\sim}{\alpha}\ \varphi(\beta, y)\}$$

where $\varphi(x, y)$ ranges over L_{PA} (in symbols $\mathrm{Ind}(\varphi, \underset{\sim}{\alpha})$; for $\varphi \in L_{\mathrm{PA}}$).

$\mathrm{I}(<\alpha)$ equals the sum of $\mathrm{I}(\beta)$: for $\beta < \alpha$.

Gentzen [7] proved $\mathrm{I}(\varepsilon_{\alpha}) \equiv \mathrm{I}(<\varepsilon_{\alpha+1})$ under the assumption that an ε-system for $\varepsilon_{\alpha+1}$ is defined. His systems are different, but this theorem can as well be proved for ε-systems of the present paper. It follows that if an ε-system for α is given then $\mathrm{I}(\alpha) \equiv \mathrm{I}(\varepsilon_{\beta})$ where ε_{β} is the greatest epsilon ordinal $\leqslant \alpha$.

Assume now that $\lambda(x, y)$ together with the accompanying formulas defines an ε-system of notation in $I\Sigma_1$.

1.6. Lemma ($I\Sigma_1$). (1) *For each* α *with* $0<\alpha<\lambda$ *there exist* $\alpha_0, \ldots, \alpha_n$ *such that* $\alpha=\omega^{\alpha_0}\dot{+}\cdots\dot{+}\omega^{\alpha_n}$. *This representation (called the Cantor normal form) is unique. Moreover,* $\alpha_0\geqslant\alpha_1\geqslant\cdots\geqslant\alpha_n$.

(2) *Assume that* $\alpha=\omega^{\alpha_0}\dot{+}\cdots\dot{+}\omega^{\alpha_n}$, $\beta=\omega^{\beta_0}\dot{+}\cdots\dot{+}\omega^{\beta_m}$. *Then* $\alpha<\beta\Leftrightarrow$ $[(a)\vee(b)]$, where

(a) $n<m\wedge\forall i\leqslant n\ \alpha_i=\beta_i$,

(b) $\exists j\leqslant n, m\ [\forall i<j\ (\alpha_i=\beta_i)\wedge\alpha_j<\beta_j]$.

(3) *Assume that* $\alpha\dot{+}\beta\downarrow$. *Then*

$$\gamma<\alpha\dot{+}\beta\ \Leftrightarrow\ [\gamma<\alpha\vee\exists\delta<\beta\ \gamma=\alpha\dot{+}\delta],\qquad \delta<\beta\rightarrow\alpha\dot{+}\delta\downarrow.$$

(4) *Assume that* $\omega^\beta\downarrow$. *Then*

$$\beta\in Lim\ \rightarrow\ \omega^\beta=\lim_{\gamma<\beta}\omega^\gamma,\qquad \beta=\gamma+1\ \rightarrow\ \omega^\beta=\lim_n\omega^\gamma n.$$

(5) *Assume that* $\varepsilon_\beta\downarrow$. *Then*

$$\beta\in Lim\ \rightarrow\ \varepsilon_\beta=\lim_{\gamma<\beta}\varepsilon_\gamma,\qquad \beta=\gamma+1\ \rightarrow\ \varepsilon_\beta=\lim_n\omega_n^{\varepsilon_\gamma+1},$$

where in general ω_n^α *is the obvious iteration of the* ω^α, $\omega_0^\alpha=\alpha$.

Proof. We work in $I\Sigma_1$. (1) is a direct consequence of Definition 1.4(iii)–(v).

We now prove (2).

It is enough to show the implication $[(a)\vee(b)]\rightarrow\alpha<\beta$ since the condition (a) $\vee$ (b) defines a linear ordering on the set of all finite increasing sequences.

If (a) holds, then there exists $\delta\neq0$ such that $\beta=\alpha\dot{+}\delta$, whence by (vii) it follows that $\alpha<\beta$. Assume now that (b) is valid. Then there exist γ, δ_1, δ_2 and $j\leqslant m$, n such that $\alpha=\gamma\dot{+}\omega^{\alpha_j}\dot{+}\delta_1$, $\beta=\gamma\dot{+}\omega^{\beta_j}\dot{+}\delta_2$ and $\alpha_j<\beta_j$. By (viii) it follows that $\omega^{\alpha_j}\dot{+}\delta_1<\omega^{\beta_j}$. Hence by (vii) we have $\omega^{\alpha_j}\dot{+}\delta_1<\omega^{\beta_j}\dot{+}\delta_2$ and $\gamma\dot{+}\omega^{\alpha_j}\dot{+}\delta_1<\gamma\dot{+}\omega^{\beta_j}\dot{+}\delta_2$, which was to be proved.

(3) follows immediately from (1) and (2).

To show (4), assume that $\omega^\beta\downarrow$. By (vii), $\omega^\gamma n<\omega^\beta$ for all n, if $\gamma<\beta$. Hence the proof of (4) reduces to the proof of the equality $\omega^\beta=\sup\{\omega^\gamma n\colon\gamma<\beta,\ n$ arbitrary$\}$. Assume that $\alpha<\omega^\beta$. It is enough to show that α is not an upper bound of the set $\{\omega^\gamma n\colon\gamma<\beta,\ n$ arbitrary$\}$.

By (iv) there exist γ, δ such that $\alpha=\omega^\gamma\dot{+}\delta$. By (viii), $\gamma<\beta$, but by (1), (2) there exists n such that $\alpha=\omega^\gamma\dot{+}\delta<\omega^\gamma n$.

To prove (5) assume that $\varepsilon_\beta\downarrow$. Since $\omega_n^{\varepsilon_\gamma+1}<\varepsilon_{\gamma+1}$ for all n (this follows from (i), (viii) and (ix)), the proof of (5) reduces to proving the equality: $\varepsilon_\beta=\sup\{\omega_{n+1}^{\varepsilon_\gamma+1}\colon\gamma<\beta,\ n$ arbitrary$\}$. To do this it is enough to show the following: $\forall\alpha<\varepsilon_\beta\ \exists\gamma<\beta\ \exists n\ \alpha<\omega_{n+1}^{\varepsilon_\gamma+1}$.

We prove this by induction on $\ulcorner\alpha\urcorner$. Assume that our claim holds for all 'ordinals' such that $\ulcorner\alpha_0\urcorner < \ulcorner\alpha\urcorner$ and let $\alpha < \varepsilon_\beta$, $\alpha = \omega^\delta \dot{+} \delta_1$. If $\delta_1 = 0$ and $\delta = \varepsilon_\gamma$, then $\alpha = \varepsilon_\gamma$ and $\gamma < \beta$. Assume then that $\delta_1 \neq 0$ or δ is not of the form ε_γ. Then by (v), (vi), $\ulcorner\delta\urcorner < \ulcorner\alpha\urcorner$, and by (viii), $\delta < \varepsilon_\beta$. By inductive assumption there exist γ, n such that $\delta < \omega_n^{\varepsilon_\gamma+1}$. Hence, by (viii), $\alpha = \omega^\delta \dot{+} \delta_1 < \omega_{n+1}^{\varepsilon_\gamma+1}$, which finishes the proof. □

1.7 Note. (i) One can prove that every recursive ordinal number has an ε-system of notation in $\mathrm{I}\Sigma_1$ (of class Σ_1). To see this it is enough to observe that each basic system of notation in $\mathrm{I}\Sigma_1$ can be extended to a set of terms forming an ε-system in $\mathrm{I}\Sigma_1$ in such a way that if λ defines a well-ordering then this extension is also well-ordered.

(ii) The natural system of notation for Γ_0 built up from symbols of operations: 0, $\alpha \dot{+} \omega^\beta$, $\kappa(\alpha, \beta)$ (definition in [5]) is a system of ε-notation in $\mathrm{I}\Sigma_1$ (actually, in $\mathrm{I}\Delta_0 + \exp$ and of class $\Delta_0(2^x)$).

1.8. Definition. Assume that λ together with the accompanying formulas defines an ε-system of notation in $\mathrm{I}\Sigma_1$ (of class Σ_1). We say that $P(\alpha, n) = \beta \in \Sigma_1$ defines a B^+-*ε-system of sequences for λ in* $\mathrm{I}\Sigma_1$ if and only if P defines a B^+-system for λ in $\mathrm{I}\Sigma_1$ and the following conditions are provable in $\mathrm{I}\Sigma_1$:

(i) $(\alpha \dot{+} \beta)_n = \alpha \dot{+} \beta_n$,
(ii) $(\omega^{\alpha+1})_n = \omega^\alpha(n+1) + 1$,
(iii) $(\omega^\alpha)_n = \omega^{\alpha_n}$ if $\alpha \in Lim \wedge \alpha \neq 0$ and α is not of the form ε_γ,
(iv) $(\varepsilon_0)_n = \omega_{n+1}^1$,
(v) $(\varepsilon_{\beta+1})_n = \omega_{n+1}^{\varepsilon_\beta+1}$,

for all α, β, γ such that $\alpha \dot{+} \beta$, ω^α, $\omega^{\alpha+1}$, $\varepsilon_{\beta+1} < \lambda$.

I.2. The existence theorems and the properties of the Hardy hierarchies

The first main aim of this section is to construct B^+-systems and B^+-ε-systems of sequences in $\mathrm{I}\Sigma_1$.

The second main aim is to show that the main properties of the hierarchy H_α^P: $\alpha < \lambda$ are provable in $\mathrm{I}\Sigma_1$, provided P is a B^+-system in $\mathrm{I}\Sigma_1$.

In contrast to the construction of the B-systems in Schmidt [28] we construct B^+-systems in one step avoiding transfinite induction which is not available in $\mathrm{I}\Sigma_1$, but the general idea of the construction is similar.

2.1. Theorem. (i) *If $\lambda(x, y) \in \Sigma_1$ defines a basic system of notations in the theory $\mathrm{I}\Sigma_1$ then there exists $P(x, y) = z \in \Sigma_1$ which defines a B^+-system for λ in $\mathrm{I}\Sigma_1$.*

Proof. Observe that there exists a formula $Q(x, y) = z$ of class Δ_1 in $\mathrm{I}\Sigma_1$ which defines a system of sequences for λ in $\mathrm{I}\Sigma_1$ having the properties:

(a) $\mathrm{I}\Sigma_1 \vdash \alpha_0^Q < \beta < \alpha \rightarrow \ulcorner\alpha\urcorner < \ulcorner\beta\urcorner$,
(b) $\mathrm{I}\Sigma_1 \vdash \alpha_n^Q < \beta < \alpha \rightarrow n < \ulcorner\beta\urcorner$.

Namely, the formula

$$\alpha = \beta + 1 \vee Lim(\alpha) \wedge \beta = \max\{\gamma < \alpha\colon \ulcorner\gamma\urcorner \leqslant \ulcorner\alpha\urcorner \vee \ulcorner\gamma\urcorner \leqslant n\} + n$$

defines such a system.

We now define a B^+-system of sequences β_n for λ in $I\Sigma_1$. We define β_n informally in $I\Sigma_1$. It is enough to define β_n for all limit β. Let β_0 be the smallest number in the sequence β_k^Q: $k = 0, 1, \ldots$ satisfying the condition

(c) $\forall\alpha\, \forall m\, [\alpha_m^Q < \beta < \alpha \rightarrow \alpha_m^Q + 1 \leqslant \beta_k^Q]$

Then we define $\beta_n = \beta_{k+n}^Q$ where $\beta_0 = \beta_k^Q$.

By (a) and (b) the condition (c) is equivalent to

$$\forall \ulcorner\alpha\urcorner < \ulcorner\beta\urcorner\, \forall m < \ulcorner\beta\urcorner\, [\alpha_m^Q < \beta < \alpha \rightarrow \alpha_m^Q + 1 \leqslant \beta_k^Q]$$

of class Δ_1 in $I\Sigma_1$. Hence the sequence β_n is of class Σ_1.

It follows directly from the definition of β_0 that $I\Sigma_1 \vdash \alpha_m^Q < \beta < \alpha \rightarrow \alpha_m^Q + 1 \leqslant \beta_0$, which implies that the system β_n is a B^+-system. □

In applications of B^+-ε-systems the next theorem is important, saying that there are sufficiently many such systems.

2.2. Theorem. *If λ together with the accompanying formulas defines an ε-system of notation in $I\Sigma_1$, then there exists $P(\alpha, n) = \beta \in \Sigma_1$ which define a B^+-ε-system for λ in* $I\Sigma_1$.

Note. One can show that every recursive ordinal has an ε-system of notation in $I\Delta_0 + \exp$, say λ, which has a B^+-ε-system of sequences in $I\Delta_0 + \exp$, say P. To construct P for λ we need, however, some properties which are stronger than those listed in 1.4.

Sketch of proof of 2.2. By 2.1 there exists $Q(x, y) = z \in \Sigma_1$ which defines a B^+-system for λ in $I\Sigma_1$. Instead of writing down the formula P explicitly we define α_n^P ($= P(\alpha, n)$) informally within $I\Sigma_1$.

We define $(\varepsilon_0)_n^P = \omega_{n+1}^1$. For $\gamma \in Lim$ we define $(\beta \dotplus \varepsilon_\gamma)_n^P = \beta \dotplus \varepsilon_{\gamma_n^Q}$. Next according to 1.8(ii) and (v) we define

$$(\beta \dotplus \omega^{\gamma+1})_n^P = \beta \dotplus \omega^\gamma(n+1) + 1, \qquad (\beta \dotplus \varepsilon_{\gamma+1})_n^P = \beta \dotplus \omega_{n+1}^{\varepsilon_\gamma + 1}.$$

Before we define fundamental sequences for the remaining ordinals we show that $\alpha = \lim \alpha_n^P$ for the above numbers α (the equality $\alpha = \lim \alpha_n^P$ presupposes that $\alpha_n^P\!\downarrow$ for all n).

Let $\alpha = \beta \dotplus \varepsilon_\gamma$, $\gamma \in Lim$. By 1.4(ix), $\varepsilon_{\gamma_n^Q} < \varepsilon_\gamma$. Hence by 1.6(3), $\alpha_n^P = \beta \dotplus \varepsilon_{\gamma_n^Q}$ is defined and by 1.4(vii), $\alpha_n^P < \beta \dotplus \varepsilon_\gamma = \alpha$.

Next if $\delta < \beta \dotplus \varepsilon_\gamma$, then by 1.6(3), (5), $\delta < \beta \dotplus \varepsilon_{\gamma'}$, for some $\gamma' < \gamma$. Hence $\delta < \beta \dotplus \varepsilon_{\gamma_m^Q}$, where m is sufficiently large and we proved that $\alpha = \lim \alpha_n^P$.

The equality $\beta + \omega^{\gamma+1} = \lim_n(\beta + \omega^{\gamma_n})$ follows from 1.6(3), (4), and $\beta \dot{+} \varepsilon_{\gamma+1} = \lim_n (\beta \dot{+} \omega_n^{\varepsilon_\gamma+1})$ by 1.6(3), (5).

For the remaining numbers α, i.e. (by (1.4(i), (iv)) for the numbers of the form $\beta \dot{+} \omega^\gamma$, where $\gamma \in Lim$ and γ is not of the form ε_δ, we define α_n^P inductively on codes according to the recursive condition $(\beta \dot{+} \omega^\gamma)_n^P = \beta \dot{+} \omega^{\gamma_n^P}$. The inductive character of this condition is a consequence of the inequality $\ulcorner\gamma\urcorner < \ulcorner\beta \dot{+} \omega^\gamma\urcorner$, which follows from 1.4(v), (vi). The equality $\alpha = \lim_n \alpha_n^P$ follows from $\gamma = \lim_n \gamma_n^P$ by 1.6(3), (4).

Hence by induction on $\ulcorner\alpha\urcorner$ we show that $\forall\alpha < \lambda\ \alpha = \lim_n \alpha_n^P$.

To show that P is a B^+-system of sequences we write this as $\forall\delta\ \forall\alpha\ (\alpha_n + 1 < \delta < \alpha \rightarrow \alpha_n + 1 \leqslant (\delta)_0)$. We use induction on $\ulcorner\delta\urcorner$. We omit further part of the proof. It is just a rather mechanical checking of all the cases for α according to Definition 1.4. □

Before the formal definition of the hierarchy H_α^P in $\mathrm{I}\Sigma_1$ we give formal definitions of the αth iterate f_α^P in the Hardy style for (codes of) partial functions $f: \subseteq \omega \rightarrow \omega$ with finite domains and study this notion. Let P be the basic system of notations.

We look for a formula $y = f_\alpha^P(x)$ of class Σ_1 defining a partial function in $\mathrm{I}\Sigma_1$ depending on f, α, x for which the following conditions are provable in $\mathrm{I}\Sigma_1$:

(i) $f_0^P(x) = x$ for all x,

(ii) $f_{\alpha+1}^P(x) \simeq f_\alpha^P(f(x))$ for all $x \in \operatorname{dom} f$, $\alpha < \lambda$,

(iii) $f_\alpha^P(x) \simeq f_{\alpha_x^P}^P(x)$ for $x \in \operatorname{dom} f$, $\alpha \in \lambda \cap Lim$,

(iv) $f_\alpha^P(x)\downarrow \rightarrow x \in \operatorname{dom} f$.

There exist many Σ_1-formulas satisfying conditions (i)–(iv) but there is no reason to suppose that any two formulas satisfying these conditions are equivalent in $\mathrm{I}\Sigma_1$. Therefore we fix one most natural formula.

2.3. Definition. $y = f_\alpha^P(x)$ is a formula which is equivalent in $\mathrm{I}\Sigma_1$ to the following statement: there exist finite sequences $\langle\alpha^i : i \leqslant l\rangle$, $\langle y^i : i \leqslant l\rangle$ such that $\alpha^0 = \alpha$, $y^0 = x$, $\alpha^l = 0$, $y^l = y$ and that for each $i < l$, $\alpha^{i+1} = (\alpha^i)_{y^i}^P$ and $y^{i+1} = f(y^i)$ when $\alpha^i \notin Lim$, or $y^{i+1} = y^i$ when $\alpha^i \in Lim$. (It follows that $y = f_{\alpha_i}^P(y^i)$ and that $f_\alpha^P(x) \simeq f^k(x)$ for some $k \leqslant l$.)

To define the αth iterate of the infinite function we use the following lemma. Let us define $f \mid [a, b] = f \cap [a, b]^2$.

2.4. Lemma ($\mathrm{I}\Sigma_1$). *If* $\forall x \in \operatorname{dom} f\ x \leqslant f(x)$ *then for every* a, b

(1) $\forall n\, f^n \mid [a, b] = (f \mid [a, b])^n$.

(2) $\forall\alpha < \lambda\, f_\alpha^P \mid [a, b] = (f \mid [a, b])_\alpha^P$.

Proof. (1) The inclusion $f^n \mid [a, b] \supseteq (f \mid [a, b])^n$ is obvious. To prove the opposite inclusion assume that $x, f^n(x) \in [a, b]$. Since $x \leqslant f(x) \leqslant \cdots \leqslant f^n(x)$, also $f(x), \ldots, f^{n-1}(x) \in [a, b]$. This implies by induction that $(f \mid [a, b])^i = f^i(x)$ for $i = 1, \ldots, n$ and the point is proved.

(2) Obviously $f^P_\alpha \mid [a, b] \supseteq (f \mid [a, b])^P_\alpha$. Assume now that $x, f^P_\alpha(x) \in [a, b]$. The witness sequence $\langle y^i : y < l\rangle$ from Definition 2.3 ($y^0 = x$, $y^l = f_\alpha(x)$) is composed of the iterations $x, f(x), \ldots, f^k(x)$ for some $k \leqslant l$. By (1), $f^j(x) = (f \mid [a, b])^j(x)$ for all $j \leqslant k$. Hence $f^P_\alpha(x) = (f \mid [a, b])^P_\alpha(x)$ and the point is proved. □

The correctness of the following definition is an immediate consequence of 2.4.

2.5. Definition ($\mathrm{I}\Sigma_1$). Let G denote a function of arbitrary class (not necessary total) such that $\mathrm{I}\Sigma_1 \vdash \forall x\, (x \leqslant G(x))$. Then the αth iterate of G, in the Hardy style, G^P_α denotes the function defined as follows:

$$G^P_\alpha(x) = y \Leftrightarrow \exists x\, (\text{"}G \mid [0, z] \text{ is codable"} \wedge (G \mid [0, z])^P_\alpha(x) = y),$$

Note. (1) Observe that in particular

$$\mathrm{I}\Sigma_1 \vdash G^P_\alpha(x) = y \Leftrightarrow (G \mid [0, y])^P_\alpha(x) = y).$$

(2) If G denotes a function of the class Σ_n in $\mathrm{I}\Sigma_n$, then all parts $G \cap [0, z]^2$ are codable and $G^P_\alpha(x) = y$ is of the class Σ_n in $\mathrm{I}\Sigma_n$. In particular the αth iterate of the successor function is well defined in $\mathrm{I}\Sigma_1$ and is of the class Σ_1: we shall denote it H^P_α: $\alpha < \lambda$. We stress that $\mathrm{I}\Sigma_1$ proves "H^P_α is total" only for $\alpha < \omega^\omega$ in the case when P is B-ϵ-system (because H_{ω^ω} majorizes all primitive recursive functions).

We now investigate the properties of the hierarchy f^P_α defined with the use of a B^+-system of sequences P. We write simply f_α for f^P_α, and α_n for α^P_n.

To formulate the main technical lemma we need the following notions. We write $\alpha \rightarrow_x \beta$ if there exists a finite sequence $\alpha = \alpha^0, \alpha^1, \ldots, \alpha^l = \beta$ such that $(\alpha^i)^P_x = \alpha^{i+1}$ for $0 \leqslant i < l$. Instead of $\alpha \rightarrow_0 \beta$ we also write $\alpha \rightarrow \beta$. This lemma will be used to a reduction of transfinite induction to finite induction, called a finitization of transfinite induction.

2.6. Lemma ($\mathrm{I}\Sigma_1$; on finitization). *For every finite partial function $f \neq \emptyset$ such that* $\forall x \in \operatorname{dom} f \;\; 0 < x < f(x)$,

(i) *the set $X_f = \{\alpha : \exists x \in \operatorname{dom} f \;\; f^P_\alpha(x)\downarrow\}$ is finite,*

(ii) $f_\alpha(x)\downarrow \wedge\, y \leqslant x \rightarrow \exists z \geqslant x\, f_{\alpha_y}(z)\downarrow$,

(iii) $f_\alpha(x)\downarrow \rightarrow \alpha \rightarrow_x 0$.

Proof. To prove (ii) it is enough to show the following implication:

$$\alpha > \beta \geqslant \alpha_n \wedge f_\beta(x)\downarrow \;\rightarrow\; \exists y \geqslant x\, f_{\alpha_n}(y)\downarrow.$$

For the proof, assume that $\alpha > \beta \geqslant \alpha_n$ and $f_\beta(x)\downarrow$. Let $\beta^0, \beta^1, \ldots, \beta^l$ be the sequence which is a witness for $z = f_\beta(x)\downarrow$ (see Definition 2.3). Let $i \leqslant l$ be the greatest number such that $\beta^i \geq \alpha_n$. Since there exists m such that $\beta^{i+1} = (\beta^i)_m$ it follows that $\beta^i = \alpha_n$, because P is a B-system (in fact a B$^+$-system) and otherwise we would have $\beta^{i+1} \geq \alpha_n$ which is a contradiction.

By the definition of $z = f_\beta(x)$, we have $z = f_{\beta^i}(y)$ for a certain $y \geq x$, which finishes the proof of (ii).

Hence, notice that (iii) is an immediate consequence of (i) and (ii). Hence it remains to show (i).

We fix $a \geq 1$. Let $\Phi(b)$ denote the following hypothesis: for every $f :\subseteq [b, a] \to [b, a]$, $f \neq \emptyset$, such that $\forall x \in \operatorname{dom} f\ x < f(x)$, the set X_f is finite.

Hence it remains to show that $\forall b \in [1, a]\ \Phi(b)$.

We use downward induction on b. $\Phi(a)$ is vacuously true. Now assume $\Phi(b+1)$, where $b \in [1, a-1]$. To prove $\Phi(b)$, let $f :\subseteq [b, a] \to [b, a]$, $f \neq \emptyset$ be such that $\forall x \in \operatorname{dom} f\ x < f(x)$. We define $g = f \restriction [b+1; a]$; obviously $g :\subseteq [b+1, a] \to [b+1, a]$ and let $\alpha = \beta + 1 \in X'_f = \{\alpha \notin Lim : f_\alpha(b)\downarrow\}$. Hence $f_\alpha(b) = f_\beta(f(b)) = g_\beta(f(b))$ and $g_\beta(f(b))\downarrow$, which implies that $\beta \in X_g$. Thus X'_f is finite. Since $\{\alpha \notin Lim : \alpha \in X_f\}$ is included in $X'_f \cup X_g$, it is also finite.

Let $\alpha \in Lim$, $\alpha \in X_f$ and let $\alpha^0, \alpha^1, \ldots, \alpha^l = 0$ be the witness sequence for $z = f_\alpha(b)$. Since $\alpha^l = (\alpha^{l-1})_m$ for some $m \geq b$, $\alpha^{l-1} = 1$. Let $k \leq l$ be the smallest natural number such that $\alpha^{k+1} \notin Lim$. We let $\alpha^{k+1} = \beta + 1$.

Obviously $\alpha = \alpha^0 > \alpha^1 = \alpha_b \geq \alpha_{b-1} + 1$, and moreover $\alpha^1, \ldots, \alpha^k, \alpha^{k+1}$ are no smaller than $\alpha_{b-1} + 1$, which we now show by induction. Assume that $\alpha^i \geq \alpha_{b-1} + 1$ for some i satisfying $1 \leq i < k+1$. Hence $\alpha > \alpha^i > \alpha_{b-1} + 1$, since $\alpha^i \in Lim$ and $\alpha = \alpha^0$. Thus using the fact that P is a B$^+$-system it follows that $\alpha^{i+1} = (\alpha^i)_b \geq \alpha_{b-1} + 1$.

Finally, we infer that $\beta \geq \alpha_{b-1}$.

Since $\beta + 1 \in X_f$, we have $\beta \in X_g$ and hence by (ii) also $\alpha_{b-1} \in X_g$. We have proved that

$$\{\alpha \in Lim : \alpha \in X_f\} \subseteq X_g \cup \{\alpha \in Lim : \alpha_{b-1} \in X_g\}.$$

To conclude that the set considered is finite it is enough to show the following implication: $\alpha_{b-1} = \gamma_{b-1} \wedge \alpha, \gamma \in Lim \to \alpha = \gamma$.

Suppose that the assumption is satisfied and let $\alpha < \gamma$. Then also $\gamma_{b-1} < \alpha < \gamma$. Hence $\gamma_{b-1} + 1 < \alpha < \gamma$, which implies that $\gamma_{b-1} + 1 \leq \alpha_0$, and this contradicts our assumption. □

Assume that P defines a B$^+$-system of sequences for λ and PA, as in the previous lemma. If f and g are partial then let us define

$$f(x) < g(x) \Leftrightarrow (\text{"}g(x) \text{ is undefined"} \vee g(x)\downarrow \wedge f(x) < g(x)).$$

The first assertion of the following lemma is a formalized local counterpart of the property (*) of the introduction to this chapter.

2.7. Lemma ($I\Sigma_1$). *For every finite partial and increasing function $f \neq \emptyset$ such that $\forall x \in \operatorname{dom} f\ \ 0 < x < f(x)$ and for every $\alpha < \lambda$*

(i) *$f_{\alpha_y}(x) < f_\alpha(x)$ for all $0 \leq y < x \in \operatorname{dom} f$,*

(ii) *f_α is increasing.*

Proof. It is enough to show our assertion for ordinals α belonging to $X_f = \{\alpha\colon \exists x \in \operatorname{dom} f\ f_\alpha(x)\downarrow\}$.

We use induction on the ordering $<$ on λ restricted to X_f, which by 2.6 is finite. And hence our induction is the usual mathematical induction.

Take $\alpha \in X_f$ and assume that the lemma is true for $\beta < \alpha$, $\beta \in X_f$. We divide the inductive steps into two sorts, for $a \notin Lim$ and for $\alpha \in Lim$, as in the proof by transfinite induction.

Case 1: $\alpha = \beta + 1$. Then $f_\alpha(x) \simeq f_\beta(f(x))$ for $x \in \operatorname{dom} f$. In particular, we have $\beta \in X_f$. Since $x < f(x)$ for $x \in \operatorname{dom} f$ and f_β is increasing, $f_\beta(x) < f_\alpha(x)$, which proves (i), because $\alpha_y = \beta$. Since f is increasing, $f_\beta \circ f$ is also increasing, which proves (ii).

Case 2: $\alpha \in Lim$. For (i) take $y < x$, $x \in \operatorname{dom} f$ and assume that $f_\alpha(x)\downarrow$. Hence $\alpha_x \in X_f$ and by 2.6(iii), $\alpha_x \to 0$. Since P is a B-system of sequences, it follows that $\alpha_x \to \alpha_y$. Take the sequence $\gamma^0, \gamma^1, \ldots, \gamma^l$ defined by $\gamma^0 = \alpha_x$, $\gamma^l = \alpha_y$ and $\forall i < l\ \gamma^{i+1} = (\gamma^i)_0$. By 2.6(ii), $\gamma^i \in X_f$ for $i \leq l$. Hence we infer by the inductive assumption that $f_{\gamma^{i+1}}(x) < f_{\gamma^i}(x)$ for $i < l$. Therefore $f_{\alpha_y}(x) < f_{\alpha_x}(x)$.

For (ii), take $y < x$, $y, x \in \operatorname{dom} f$. It follows that $f_{\alpha_y}(x) < f_{\alpha_x}(x)$. Since, by the inductive assumption, f_{α_y} is increasing, also $f_{\alpha_y}(y) < f_{\alpha_x}(x)$, i.e., $f_\alpha(y) < f_\alpha(x)$, which finishes the proof. □

Corollary 1. $I\Sigma_1 \vdash$ "*$H^P_{\alpha_y}(x) < H^P_\alpha(x)$ for all $y < x \in \mathbb{N}$, $\alpha < \lambda$*".

Proof. Indeed, assume that $0 \leq y < x$ and denote $z = H_\alpha(x)$ (if $H_\alpha(x)$ is undefined then there is nothing to prove). Then $z = (H \mid [0, z])_\alpha(x) > (H \mid [0, z])_{\alpha_y}(x)$, by 2.7(i). Since $(H \mid [0, z])_{\alpha_y}(x)$ is defined, it equals $H_{\alpha_y}(x)$ and the corollary is proved. □

Corollary 2. $I\Sigma_1 \vdash$ "*for each $\alpha < \lambda$, H_α is increasing*".

II. Logic and combinatorics

In this chapter we introduce and study the so-called a-skeletons, the basic semantic notion in this paper. Roughly speaking, a-skeletons are finite sets of diagonally indiscernible elements for the family of Δ_0-formulas obtained by bounding all quantifiers in formulas less than a.

In Sections 1 and 2, we study the logical properties of a-skeletons. In many combinatorial constructions having a logical goal we can use only a-skeletons instead of classical models; as an example, we show how this can be done in the proof of the theorem of Paris–Harrington [21] on the independence of a combinatorial principle from PA. (This new proof has some common features with the proof of independence presented by Kurata [16].)

This gives a basic example of the combinatorial-logical method of the proof, which was generally described in the Introduction. This method will be fully described and developed in the next chapters. It enables one to make α-fold iterations of combinatorial constructions without transfinite induction up to α, which is used to study the proof-theoretical strength of the theories $\mathrm{I}(\varepsilon_\alpha)$ and $\mathcal{R}_\alpha(\mathrm{PA})$.

In Sections 2 and 3, we examine the problem of the existence of a-skeletons having some additional properties. We generalize and strengthen the construction of diagonally indiscernible elements of [21]. Moreover, we examine the question of how large a set A should be in order that there exists a β-large (in the sense of Solovay and Ketonen [12]) a-skeleton B such that $B \subseteq A$. These results are the next preparatory step in the combinatorial-logical studies of the theories $\mathrm{I}(\varepsilon_\alpha)$ and $\mathcal{R}_\alpha(\mathrm{PA})$.

II. 1. Logical properties of a-skeletons

Before we define the notion of a-skeleton, we define precisely the notion of a finite set of diagonally indiscernible elements.

1.1. Definition ($\mathrm{I}\Delta_0 + \exp$). Assume that $\mathcal{R}(\bar{b}, \bar{c}) \in \Delta_0(2^x)$ is a binary relation defined on pairs $(\bar{b}, \bar{c})$ of sequences of length m and n respectively. We say that a finite set A is a *set of diagonally indiscernible elements with respect to* $\mathcal{R}$ iff

$$\forall a \in A\, \forall \bar{b} < a\, \forall \bar{c}, \bar{d}\, \{\bar{c}, \bar{d} \in [A \setminus [0, a]]^n \to \mathcal{R}(\bar{b}, \bar{c}) \Leftrightarrow \mathcal{R}(\bar{b}, \bar{d})\}.$$

Here the inequality $\bar{b} < a$ means that all terms of the sequence $\bar{b}$ are $<a$. If $m = 0$, then the quantifier $\forall \bar{b} < a$ is superfluous. Observe that then A is a set of diagonally indiscernible elements for $\mathcal{R}(\bar{c})$ (in the above sense) iff $A \setminus \{\min A\}$ is a set of indiscernible elements for $\mathcal{R}(\bar{c})$, in the usual sense.

1.2. Definition ($\mathrm{I}\Delta_0 + \exp$). Let $\theta^*(\bar{x}, \bar{y})$ with $\bar{y} = y_1, \ldots, y_n$ denote a formula of class Δ_0 obtained from $\theta(\bar{x}) \in L_{\mathrm{PA}}$ by bounding the quantifiers of the deepest uniform blocks by y_n, the subsequent blocks by y_{n-1} and so on; the last one will be bounded by y_1. The number n is called the arithmetical range of the formula θ. We then write $\mathrm{a.r.}(\theta) = n$. For example, $(\exists \bar{x}\, (\theta \vee \eta))^* := \exists \bar{x} < y_{n+1}\, (\theta^* \vee \eta^*)$, where $n = \max(\mathrm{a.r.}(\theta), \mathrm{a.r.}(\eta))$.

Let $\mathrm{Tr}_{\Delta_0} \in \Delta_0(2^x)$ denote the truth predicate for Δ_0-formulas. We say that A is a set of diagonally indiscernible elements for $\theta^*(\bar{x}, \bar{y})$ iff A is a set of diagonally indiscernible elements for $\mathrm{Tr}_{\Delta_0}(\theta^*(\bar{b}, \bar{c}))$.

For example $(\forall x_1\, \exists x_2\, 2x_1 < x_2)^* := \forall x_1 \leq y_1\, \exists x_2 \leq y_2\, 2x_1 < x_2 \equiv 2y_1 < y_2$.

1.3. Notation. In this paper symbols θ, η etc. denote distinguished variables of the "standard", i.e. real, language of L_{PA} which are used to denote formulas of L_{PA} in theories extending $\mathrm{I}\Delta_0+\exp$. Formulas belonging to the standard L_{PA} are denoted by φ, ψ, etc. and their Gödel numbers by $\ulcorner\varphi\urcorner$, $\ulcorner\psi\urcorner$.

Throughout the paper we also use the following convention: if A is a finite set of natural numbers, then the successive elements of A, in increasing order, are denoted by $a_0, a_1, \ldots, a_i, \ldots$; similarly for the elements of B, C and of A^k, B^k, C^k. In the theory $\mathrm{I}\Delta_0+\exp$ we identify formulas of L_{PA} with their codes. Hence instead of saying that the code of the formula η is less than a we can write simply $\eta<a$. Now we define the notion of a-skeleton.

1.4. Definition ($\mathrm{I}\Delta_0+\exp$). We say that a finite set A is an *a-skeleton* iff

(i) $|A|>a_0\geqslant a$,

(ii) A is a set of diagonally indiscernible elements for θ^* for all θ of the form $\forall\bar{x}\,\eta$, where $\eta<a$ and $\bar{x}$ is an arbitrary sequence of free variable of the formula η,

(iii) A is a set of indiscernible elements for θ^* for all θ of the form as in (ii) which are sentences.

Before we turn to the investigations of properties of a-skeletons we recall some facts from [21], which provide a motivation.

Let PH denote the following principle:

$$\forall e, k, r\, \exists M\, \forall F:[M]^e \to r\ \exists A \text{ "}A\text{ is homogeneous for } F \text{ and } |A|>a_0, k\text{"}.$$

It was proved in [21] that $\mathrm{PA}\nvdash\mathrm{PH}$. In fact, Paris and Harrington proved a stronger theorem. We denote by $\mathrm{R}(\mathrm{PA};\Sigma_1)$ the following sentence of L_{PA}:

$$\forall\theta\in\Sigma_1\,[\mathrm{PA}\vdash\theta\to\mathrm{Tr}_{\Sigma_1}(\theta)].$$

1.5. Theorem [21]. $\mathrm{PA}\vdash(\mathrm{PH}\to\mathrm{R}(\mathrm{PA};\Sigma_1))$.

Now we sketch the proof. Working informally in $\mathrm{PA}+\mathrm{PH}$ Paris and Harrington show that for every finite sequence $\theta_1,\ldots,\theta_l$ of formulas in the prenex normal form there exists a set A of diagonally indiscernible elements for $\theta_1^*,\ldots,\theta_l^*$ such that $|A|>a_0$ and $\forall a, b\in A\,(a<b\to a^2<b)$. It follows by compactness that there exists a model $M\vDash\mathrm{I}\Delta_0+\exp+\{\eta:\mathrm{Tr}_{\Pi_1}(\eta)\}$ and an infinite set $A\in M$ which is diagonally indiscernible for all θ^* where $\theta\in L_{\mathrm{PA}}$. Then they construct a suitable cut I of M which is a model for $\mathrm{PA}+\{\eta:\mathrm{Tr}_{\Pi_1}(\eta)\}$; this finishes the proof. This idea is based on the following observation, which we formulate as a separate proposition.

1.6. Proposition (PA). *Assume that $M\vDash\mathrm{I}\Delta_0+\exp$, $\varphi(x)\in L_{\mathrm{PA}}$ and $\varphi_1,\ldots,\varphi_l=\varphi$ are all subformulas of φ and that $M\vDash$"A is a set of diagonally indiscernible elements for $\varphi_1^*,\ldots,\varphi_l^*$". Moreover, assume that $M\vDash A=\{a_0,\ldots,a_k\}\wedge\forall i<k\, a_i^2<a_{i+1}$ and that A is infinite. Let J be an initial cut in $[0,k]$ closed under successors.*

Then for $I = \{a\colon \exists i \in J\ a \leq a_i\}$, *the following holds:*

$$\forall x \in I\,[(I \models \varphi(x)) \Leftrightarrow M \models \mathrm{Tr}_0(\varphi^*(x, a_{k-n+1}, \ldots, a_{k-1}, a_k)),$$

where $n = \mathrm{a.r.}(\varphi)$.

(From the above equivalence it follows that $I \models \mathrm{Ind}(\varphi)$.)

For convenience we will constantly use the following notation.

1.7. Notation ($\mathrm{I}\Delta_0 + \exp$). Assume that $A = \{a_0, \ldots, a_k\}$. We will write $A \models \theta(\bar{x})$ instead of

$$\mathrm{Tr}_{\Delta_0}(\theta^*(\bar{x}, a_{k-n+1}, \ldots, a_{k-1}, a_k)) \quad \text{for } n = \mathrm{a.r.}(\theta) \leq |A|.$$

In particular observe that for $\theta \in \Delta_0$ and for every $\bar{x} \leq a_{k-n+1}$, $A \models \theta(\bar{x})$ iff $\mathrm{Tr}_{\Delta_0}(\theta(\bar{x}))$, because then all bounds in θ are smaller than the additional bounds in θ^*.

Proposition 1.6 shows that the relation $A \models \theta(\bar{x})$ is very similar to the usual satisfaction relation for standard θ and appropriate sets A. We show that the relation $A \models \theta(\bar{x})$ defined in 1.7 when restricted to a-skeletons determines the natural (in a sense) semantics for formulas less than a (standard and nonstandard), in accordance with provability limited by a. Then we will be ready to present an alternative proof of the Paris–Harrington theorem.

Let us now comment for a moment on the problem of existence of a-skeletons. An easy strengthening of the construction of diagonally indiscernible elements from [21] gives the following result.

1.8. Lemma. $\mathrm{I}\Delta_0 + \exp + \mathrm{PH} \vdash \forall \mathrm{a}\, \exists \mathrm{A}$ "*A is an a-skeleton*" $\wedge\ \forall a, b \in A\ [a < b \rightarrow (a+1)^2 < b]$.

In the Section II.3 we prove a more general result on the existence of skeletons.

Now we study the properties of the relation $A \models \theta$. For this relation the following counterparts of the Tarski conditions for truth hold.

1.9. Proposition ($\mathrm{I}\Delta_0 + \exp$). *Assume that* A *is an a-skeleton and let* $A = \{a_0, \ldots, a_k\}$. *Then*

(i) $\theta, \neg\theta < a \rightarrow A \models \neg\theta(\bar{x}) \Leftrightarrow \neg A \models \theta(\bar{x})$,

(ii) $\theta \vee \eta < a \rightarrow A \models (\theta \vee \eta)(\bar{x}) \Leftrightarrow (A \models \theta(\bar{x}) \vee \models \eta(\bar{x}))$.

(iii) $\theta(\bar{x}, \bar{y}) < a \rightarrow \forall \bar{x} < a_{k-(r+1)}\,[A \models \forall \bar{y}\, \theta(\bar{x}, \bar{y}) \Leftrightarrow \forall \bar{y} < a_{k-r} A \models \theta(\bar{x}, \bar{y})]$, *where* r *is a number such that either* $\mathrm{a.r.}(\theta) \leq r \leq k-1$, *or* $r = k$ *and* $\bar{x}$ *is empty.*

In particular, if $\theta(\bar{x}) < a$, then $A \models \forall \bar{x}\, \theta(\bar{x}) \Leftrightarrow \forall \bar{x} < a_0\, A \models \theta(\bar{x})$, if we assume additionally that formulas are defined (coded) in such a way that $\mathrm{a.r.}(\theta) \leq \theta$. This is our standing assumption throughout the paper, as is the assumption that for all η, θ, if η is a subformula of θ, then $\eta < \theta$.

Note. In (iii) we use the convention, constantly valid in this paper, that the same variable symbol in the same formula can denote variables of different types, dependent on the context in which it appears.

For example in (iii): the symbol $\bar{x}$ in $\theta(\bar{x}, \bar{y}) < a$ denotes a metavariable which denotes a sequence of formal variables; in $A \models \forall \bar{y}\, \theta(\bar{x}, \bar{y})$ the same symbol denotes a metavariable denoting the sequence of valuations of the formal variables of the previous sequence $\bar{x}$. The symbol $\bar{y}$ in the latter context denotes a formal sequence of variables.

Proof of 1.9. (i) and (ii) are immediate consequences of 1.7. We only prove (iii) under the assumption that θ does not begin with $\forall$. The opposite case is left to the reader. Set $m = \text{a.r.}(\theta)$. Assume that $\theta(\bar{x}, \bar{y}) < a$ and $m \leq r \leq k-1$. Take $\bar{x} < a_{k-(r+1)}$.

Assume that $A \models \forall \bar{y}\, \theta(\bar{x}, \bar{y})$. Hence by 1.7, $\forall \bar{y} < a_{k-m}\, A \models \theta(\bar{x}, \bar{y})$, which implies that $\forall \bar{y} < a_{k-r}\, A \models \theta(\bar{x}, \bar{y})$. Now assume that the latter sentence is true. Hence

$$\forall \bar{y} < a_{k-r} \operatorname{Tr}_0(\theta^*(\bar{x}, \bar{y}; a_{k-m+1}, \ldots, a_{k-1}, a_k)),$$

which implies that $\operatorname{Tr}_0((\forall \bar{y}\, \theta)^*(\bar{x}, a_{k-m+1}, \ldots, a_{k-1}, a_k))$. Since the elements of A are diagonally indiscernible for $(\forall \bar{y}\, \theta)^*$, we have

$$\operatorname{Tr}_0((\forall \bar{y}\, \theta)^*(\bar{x}, a_{k-m}, a_{k-m+1}, \ldots, a_{k-1}, a_k)), \quad \text{i.e.} \quad A \models \forall \bar{y}\, \theta(\bar{x}). \qquad \square$$

Now we need a few simple definitions in PA. Let Φ be a definable set of sentences of L_{PA}. We assume that the relation $\Phi \vdash \theta$ is defined on the base of the Hilbert type of proof. We assume that our formal system of classical logic has only two rules of inference: modus ponens and generalization, and only two axiom schemes concerning quantifiers:

$\forall x\, (\eta \rightarrow \theta) \rightarrow (\eta \rightarrow \forall x\, \theta)$, where x is not free in η;

$\forall x\, \eta \rightarrow \eta(t)$, where t ranges only over simple terms of the form: 0, 1, x, $x+1$, $x+y$, $x \cdot y$ with the well-known restriction on them.

If $p = (\theta_0, \ldots, \theta_k)$ is a proof then the number $\max_{i \leq k} \theta_i$ will be called the width of p.

We write $\Phi \vdash_d \theta$ if the formula θ has a proof within Φ with width less than d. Now assume that the set A is an a-skeleton. We write $A \models_a \theta$ if $A \models \theta$ and $\theta < a$, and $A \models_a \Phi$ if $\forall \theta\, [\theta \in \Phi \wedge \theta < a \rightarrow A \models \theta]$. We say that A is $(x+1)^2$-scattered iff $\forall a, b \in A\, [a < b \rightarrow (a+1)^2 < b]$.

1.10. Lemma (PA; soundness of $\Phi \vdash_a \theta$). *Assume that A is an $(x+1)^2$-scattered a-skeleton. Then for each formula $\theta(\bar{x})$ the following implication is true:*

$$A \models_a \Phi \rightarrow (\Phi \vdash_a \theta \rightarrow A \models \forall \bar{x}\, \theta).$$

In fact this lemma is provable in $\mathrm{I}\Delta_0 + \exp$ if we assume that all basic syntactical notions are formalized in $\mathrm{I}\Delta_0 + \exp$ and if Φ is of class $\Delta_0(2^x)$.

Proof. Assume that $A \vDash_a \Phi$, $\Phi \vdash_a \theta$ and that $A = \{a_0, \ldots, a_k\}$. Let $\theta_0(\bar{x}_0), \ldots, \theta_l(\bar{x}_l)$ be a proof for θ within Φ such that $\forall i \leq l\ \theta_i < a$. We show by induction on i that $\forall i \leq l\ A \vDash \forall \bar{x}_i\ \theta_i(\bar{x}_i)$. Because of 1.9(iii) it is enough to show that for each $i \leq l$, $\forall \bar{x}_i < a_0\ A \vDash \theta_i(\bar{x}_i)$.

If $\theta_i(\bar{x}_i)$ is a substitution of a propositional tautology then the inductive conclusion is a direct consequence of 1.9(i), (ii). Now let us check the logical axioms dealing with quantifiers.

Case (a): $\theta_i(\bar{x}_i) := \forall x\,(\eta_1 \rightarrow \eta_2) \rightarrow (\eta_1 \rightarrow \forall x\ \eta_2)$. Assume that $\bar{x}_i < a_0$ and $A \vDash \forall x\,(\eta_1 \rightarrow \eta_2)$. Since $\text{a.r.}(\eta_1 \rightarrow \eta_2) \leq (\eta_1 \rightarrow \eta_2) < a_0 - 1 \leq k - 1$, we infer by 1.9(iii) that $\forall x < a_1\ A \vDash \eta_1 \rightarrow \eta_2$. Hence $\forall x < a_1\ [(A \vDash \eta_1) \rightarrow A \vDash \eta_2]$ and in consequence $(A \vDash \eta_1) \rightarrow \forall x < a_1\ A \vDash \eta_2$. Referring once more to 1.9(iii) we obtain $A \vDash \eta_1 \rightarrow \forall x\ \eta_2$. Summing up,

$$\forall \bar{x}_i < a_0\ A \vDash \forall x\,(\eta_1 \rightarrow \eta_2) \rightarrow (\eta_1 \rightarrow \forall x\ \eta_2).$$

Case (b): $\theta_i(\bar{x}_i) := \forall x\ \eta(x, \bar{x}_i) \rightarrow \eta(t, \bar{x}_i)$, where t is a simple term.

Assume that $\bar{x}_i < a_0$ and that $A \vDash \forall x\ \eta(x, \bar{x}_i)$. Obviously $\text{a.r.}(\eta) \leq \eta < a_0 - 1 \leq k - 1$. Hence $\forall x < a_1\ A \vDash \eta(x, \bar{x}_i)$. Since A is $(x+1)^2$-scattered, $t(\bar{x}_i) < a_1$ and hence $A \vDash \eta(t, \bar{x}_i)$, which finishes the proof in this case.

/ If $\theta_i(\bar{x}_i)$ belongs neither to Φ nor to the logical axioms and is a result of a generalization then the inductive step is obvious. Assume now that there exist $j, j' < i$ such that $\theta_j := \theta_{j'} \rightarrow \theta_i$. From the inductive assumptions $\forall \bar{x}_{j'} < a_0\ A \vDash \theta_{j'}(\bar{x}_{j'})$, $\forall \bar{x}_j < a_0\ A \vDash \theta_{j'}(\bar{x}_{j'}) \rightarrow \theta_i(\bar{x}_i)$ it follows obviously that $\forall \bar{x}_i < a_0\ A \vDash \theta_i(\bar{x}_i)$ and this finishes the proof. □

1.11. Note. The condition "A is $(x+1)^2$-scattered" appearing in the assumption of Lemma 1.10 can be replaced by $c_0 < a$, where c_0 is the maximum of the codes of the formulas $\forall x_1\, \exists x_2\, 2x_1 < x_2$, $\forall x_1\, \exists x_2\, (x_1+2)^2 \leq x_2$. The 'star' of the first formula is equivalent to $2y_1 \leq y_2$, and the 'star' of the second to $(y_1+1)^2 < y_2$. Since $A = \{a_0, \ldots, a_k\}$ is diagonally indiscernible for $2y_1 \leq y_2$ and $2a_0 \leq a_k$, we obtain $\forall i < k\ 2a_i \leq a_{i+1}$. Hence $2^{a_0}a_0 \leq a_k$. In particular, $(a_0+1)^2 < a_k$ and finally A is $(x+1)^2$-scattered.

1.12. Lemma ($\mathrm{I}\Delta_0 + \exp$). *If A is an a-skeleton then for every $\theta < a$ which is a particular case of the scheme of mathematical induction we have $A \vDash \theta$.*

Proof. Assume that

$$\theta := \forall \bar{y}\ [\eta(0, \bar{y}) \wedge \forall x\,(\eta(x, \bar{y}) \rightarrow \eta(x+1, \bar{y})) \rightarrow \forall x\ \eta(x, \bar{y})]$$

is less than a. Take $\bar{y} < a_0$ and assume that

$$A \vDash \eta(0, \bar{y}) \wedge \forall x\,(\eta(x, \bar{y}) \rightarrow \eta(x+1, \bar{y})).$$

By 1.9, it follows that

$$A \models \eta(0, \bar{y}) \wedge \forall x < a_1 [(A \models \eta(x, \bar{y})) \rightarrow A \models \eta(x+1, \bar{y})].$$

Thus since we have $\Delta_0(2^x)$-induction, $\forall x < a_1\, A \models \eta(x, \bar{y})$. Finally, $A \models \forall x\, \eta(x, \bar{y})$ and this finishes the proof. □

We end this section by giving a finite counterpart of the part of the presented reasoning of Paris and Harrington.

1.13. Lemma. $\mathrm{I}\Delta_0 + \exp$ *proves the following implication:*

$$\forall a\, \exists A \text{ “}A \textit{ is an } a\textit{-skeleton”} \rightarrow \mathrm{R}(\mathrm{PA}; \Sigma_1).$$

Corollary. $\mathrm{I}\Delta_0 + \exp$ *proves the following equivalencies*:

$$\mathrm{PH} \equiv \mathrm{R}(\mathrm{PA}; \Sigma_1) \equiv \forall a\, \exists A \text{ “}A \textit{ is an } a\textit{-skeleton”}.$$

Proof of 1.13. We work informally in $\mathrm{I}\Delta_0 + \exp + \forall a\, \exists A$ “A is an a-skeleton”. Assume that $\mathrm{PA} \vdash \exists x\, \eta(x)$, where $\eta(x) \in \Delta_0$. Hence there exists an $a > \exists x\, \eta(x)$ such that $\mathrm{PA} \vdash_a \exists x\, \eta(x)$. Assume moreover that a is greater than the constant c_0 from Note 1.11. Take an arbitrary a-skeleton A. Hence A is $(x+1)^2$-scattered. By 1.12, $A \models_a \mathrm{PA}$. It follows by 1.10 that $A \models_a \exists x\, \eta(x)$. Hence $\exists x < a_0\, A \models \eta(x)$. Since $\eta \in \Delta_0$, also $\exists x < a_0\, \mathrm{Tr}_0(\eta(x))$. Hence we have shown that $\mathrm{Tr}_0(\exists x\, \eta(x))$, which finishes our proof. □

1.14. Note. Another proof-theoretic proof of the above lemma follows from the results of Kurata [16]. Let T denote the theory of Paris–Harrington [21], i.e., $\mathrm{I}\Delta_0$ plus $\{c_i^2 < c_{i+1} : i \in \mathbb{N}\}$ plus the set of sentences $\{\varphi^*(c_{i_1}, \ldots, c_{i_n}) \Leftrightarrow \varphi^*(c_{j_1}, \ldots, c_{j_n}) : \varphi \in L_{\mathrm{PA}},\ i_1 < \cdots < i_n \text{ and } j_1 < \cdots < j_n\}$. Kurata considers the sentence $\mathrm{FC}_\omega(T)$ which, roughly speaking, says that every finite subset of T has a standard model. This sentence is very close to our $\forall a\, \exists A$ “A is an $(x+1)^2$-scattered a-skeleton”. He shows that $\mathrm{PA} \vdash \mathrm{FC}_\omega(T) \rightarrow R(PA, \Sigma_1)$ by interpreting in some way T in PA. He uses the following proposition.

Proposition. $\forall \varphi \in \mathrm{Sent}_{\mathrm{PA}}\, (\mathrm{PA} \vdash \varphi \rightarrow T \vdash \varphi^*(c_1, \ldots, c_n))$.

This easy proposition was not in fact proved by Kurata [16]. A non-model-theoretic proof of a very similar theorem appears in Mycielski [19]. His theorem concerns all theories, not only PA. In Lemmas 1.10 and 1.12 we point out some facts connected with width of proofs which we later extensively use.

II.2. The properties of scattered skeletons

In this section we consider some strengthening of 1.5.

2.1. Definition ($\mathrm{I}\Delta_0 + \exp$). Assume that $A = \{a_0, \ldots, a_k\}$. Then for $n \geqslant 1$, A is *Σ_n-scattered* iff

$$\forall \theta(\bar{x}, \bar{y}) \in \Pi_{n-1} [\exists \bar{y}\, \theta(\bar{x}, \bar{y}) < a_0 \rightarrow \forall i < k\, \forall \bar{a}$$
$$[\bar{a} < a_i \wedge \mathrm{Tr}_{\Sigma_n}(\exists \bar{y}\, \theta(\bar{a}, \bar{y})) \rightarrow \exists \bar{b} < a_{i+1}\, \mathrm{Tr}_{\Pi_{n-1}}(\theta(\bar{a}, \bar{b}))]].$$

We make the convention that every set A is Σ_0-scattered.

Observe that the notion "A is Σ_n-scattered" when considered in $\mathrm{I}\Sigma_n$ is of class Π_n in that theory.

Let PH_n for $n \geqslant 1$ denote the sentence

$$\forall e, k, r\, \exists M\, \forall F : [M]^e \rightarrow r\, \exists A\, [\text{"}A \text{ is homogeneous for } F\text{"}$$
$$\wedge \text{"}A \text{ is } \Sigma_{n-1}\text{-scattered"} \wedge |A| > \max(a_0, k)].$$

It is not difficult to show the following proposition.

2.2. Proposition. $\mathrm{PA} \vdash \mathrm{PH}_n \Leftrightarrow \mathrm{R}(\mathrm{PA}, \Sigma_n)$ *for* $n \geqslant 1$.

$\Leftarrow$ is standard (cf. Kurata [16]). The opposite implication can be proved model-theoretically as follows. In the first step, arguing as Paris and Harrington [21] but now in $\mathrm{PA} + \mathrm{PH}_n$ we show that

$$\forall a\, \exists A \text{ "}A \text{ is a } \Sigma_{n-1}\text{-scattered } a\text{-skeleton"}. \tag{1}$$

By compactness there exists a model $M \models \mathrm{I}\Delta_0 + \exp + \{\eta : \mathrm{Tr}_{\Pi_n}(\eta)\}$ and an infinite set $A \in M$ which is diagonally indiscernible for all θ^*, where $\theta \in L_{\mathrm{PA}}$, and which is Σ_{n-1}-scattered in M.

By Proposition 1.6 there exists a cut $I \subseteq_e \mathrm{M}$ which is a model for PA. It is also a model for $\{\eta : \mathrm{Tr}_{\Pi_n}(\eta)\}$ (which finishes the proof), because of the following proposition.

2.3. Proposition. *Let* $n \geqslant 1$. *Assume that* $M \models \mathrm{I}\Delta_0 + \exp$ *and that* $A \in M$ *and* $M \models$ *"A is Σ_n-scattered". Then for every initial cut* J *of* $[0, |A|]$ *which is closed under successors the cut* $I = \{a \in M;\ \exists i \in J\ a \leqslant a_i\}$ *is a Σ_n-elementary submodel of* M.

We omit the simple proof. Another proof of Proposition 2.2, rather proof-theoretic, follows from Kurata's results [16]. Now we present a different proof-theoretic proof. We need the following local counterpart of 2.3. Let $\theta^{\neg}$ denote the formula which is obtained from $\neg\theta$ by transporting the negation sign down to atomic subformulas.

2.4. Lemma ($I\Delta_0 + \exp$; on absoluteness). *If a set A is Σ_n-scattered and $|A| > a_0$, then for every $\theta \in \Sigma_n \cup \Pi_n$ such that θ, $\theta\neg < a_0$ and for every $\bar{b} < a_{k-m+1}$, where $k = |A| - 1$ and $m = \text{a.r.}(\theta)$, we have*

$$(A \models \theta(\bar{b})) \Leftrightarrow \mathrm{Tr}_{\Sigma_n \cup \Pi_n}(\theta(\bar{b})).$$

In particular, for all $\bar{b} < a_0$, $(A \models \theta(\bar{b})) \Leftrightarrow \mathrm{Tr}_{\Sigma_n \cup \Pi_n}(\theta(\bar{b}))$.

Proof. We can formulate this lemma in the following form:

$$\forall l \leqslant n \ I\Delta_0 + \exp \vdash \forall A \text{ “}A \text{ is } \Sigma_n\text{-scattered and } |A| > a_0 \text{ and } \forall \theta \in \Sigma_l \cup \Pi_l \text{ etc.”}$$

where the ‘etc.’ part is the same as in the original formulation.

Now we prove this lemma by metamathematical induction on $l \leqslant n$. For $l = 0$ the lemma is obvious (cf. 1.7). We make the inductive step $l \to l+1 \leqslant n$. We work in $I\Delta_0 + \exp$. To fix our ideas let $\theta \in \Sigma_{l+1} \cup \Pi_{l+1}$, $m = \text{a.r.}(\theta)$ and let $\bar{b} < a_{k-(m+1)+1} = a_{k-m}$ etc.

Case 1: $\theta := \exists \bar{x}\, \eta$, *where* $\eta \in \Pi_l$. Hence

$$A \models \theta(\bar{b}) \Leftrightarrow \exists \bar{x} < a_{k-m+1}\, A \models \eta(\bar{x}, \bar{b}).$$

By the inductive assumption this is equivalent to $\exists \bar{x} < a_{k-m+1} \mathrm{Tr}_{\Pi_l}(\eta(\bar{x}, \bar{b}))$ and obviously implies $\mathrm{Tr}_{\Sigma_n}(\theta(\bar{b}))$. The opposite implication follows because A is Σ_n-scattered.

Case 2: $\theta := \forall \bar{x}\, \eta$, *where* $\eta \in \Sigma_l$. Since $\theta\neg < a_0$, by Case 1 we have

$$A \models \theta\neg(\bar{b}) \Leftrightarrow \mathrm{Tr}_{\Sigma_n}(\theta\neg(\bar{b})).$$

It follows that

$$A \models \theta(\bar{b}) \Leftrightarrow \mathrm{Tr}_{\Pi_n}(\theta(\bar{b})). \qquad \square$$

Using 2.4 we can prove a generalization of 1.13. Now we prove a more general lemma which provides a form of reduction of the reflection principle to combinatorics. This lemma is used in the next sections. Let us recall the uniform principle of reflection in full generality. Assume that T is a theory in the language of PA, and $I\Sigma_1 \subseteq T$. Assume also that a representation of T in $I\Sigma_1$ is given. Moreover we have a primitive recursive mapping which to each number b assigns a closed term $\underline{b}$, called a term number (e.g. numeral), whose value is equal to b.

If b denotes a variable of L_{PA}, then $\underline{b}$ denotes the formalized counterpart of the term number for b such that $I\Sigma_1 \vdash$ “value of $\underline{b}$” $= b$.

2.5. Definition. Let $n \geqslant 1$.

(1) $\mathrm{R}(T; \Sigma_n)$ denotes the following sentence of L_{PA}:

$$\forall \theta \in \Sigma_n \forall b\, [T \vdash \theta(\underline{b}) \to \mathrm{Tr}_{\Sigma_n}(\theta(b))].$$

(2) $\mathrm{CR}(T; \Sigma_n)$ denotes the following sentence of L_{PA}:

$$\forall a \exists A\, [\text{“}A \text{ is a } \Sigma_{n-1}\text{-scattered } a\text{-skeleton”} \wedge A \models_a \mathrm{Ax}(T)].$$

The principle $\mathrm{CR}(T, \Sigma_n)$ is called the combinatorial reflection principle. The set $\{\mathrm{R}(T, \Sigma_n)\colon n \in \omega\}$ is denoted by $\mathcal{R}(T)$.

We have the following lemma.

2.6. Lemma. $I\Sigma_1 \vdash \mathrm{CR}(T; \Sigma_n) \rightarrow \mathrm{R}(T; \Sigma_n)$.

Proof. The proof is almost identical with the proof of 1.13.

We work informally in $\mathrm{I}\Sigma_1 + \mathrm{CR}(T; \Sigma_n)$. Assume that $T \vdash \exists y\, \eta(\underline{b}, y)$, where $\eta(x, y) \in \Pi_{n-1}$. Hence there exists an $a > \ulcorner\exists y\, \eta(\underline{b}, y), \eta(\underline{b}, y)\urcorner$ such that $\mathrm{Ax}(T) \vdash_a \exists y\, \eta(\underline{b}, y)$. Take an arbitrary Σ_{n-1}-scattered a-skeleton A such that $A \vDash_a \mathrm{Ax}(T)$. It follows by Soundness Lemma 1.10 that $A \vDash_a \exists y\, \eta(\underline{b}, y)$. Hence $\exists y < a_0\, A \vDash \eta(\underline{b}, y)$. By Absoluteness Lemma 2.4 also $\exists y < a_0 \mathrm{Tr}_{\Pi_{n-1}}(\eta(\underline{b}, y))$. Hence we have shown that $\mathrm{Tr}_{\Sigma_n}(\exists y\, \eta(b, y))$, which finishes the proof. $\square$

2.7. Since a-skeletons are models for fragments of $\mathrm{Ax}(\mathrm{PA})$, the principle $\mathrm{CR}(\mathrm{PA}; \Sigma_n)$ can be written shortly as follows:

$$\forall a\, \exists A \text{ “}A \text{ is a } \Sigma_{n-1}\text{-scattered } a\text{-skeleton”}.$$

Hence by 2.2(1), $\mathrm{I}\Sigma_1 \vdash \mathrm{PH}_n \Leftrightarrow \mathrm{CR}(PA; \Sigma_n)$.

Thus finally we obtain

Corollary. $\mathrm{I}\Sigma_1$ *proves the following equivalencies:* $\mathrm{PH}_n \Leftrightarrow \mathrm{R}(\mathrm{PA}; \Sigma_n) \Leftrightarrow \mathrm{CR}(\mathrm{PA}; \Sigma_n)$.

2.8. Note. An unsolved problem is the characterization of those theories T for which the principles $\mathrm{R}(T; \Sigma_n)$ and $\mathrm{CR}(T, \Sigma_n)$ are equivalent (over $\mathrm{I}\Sigma_1$). In Chapter III we show, among other things, that this is true for all $\mathrm{I}(\varepsilon_\alpha)$, and in Chapter IV for all $\mathcal{R}_\alpha(\mathrm{PA})$. Our method is to construct a-skeletons which are models for fragments of $\mathrm{I}(\varepsilon_\alpha)$ and of $\mathcal{R}_\alpha(\mathrm{PA})$.

Sometimes in these constructions it is essential to establish whether a given skeleton is scattered enough; cf. 1.11 and its applications in the proof of 1.13. The proposition included in Note 1.11 can be generalized. We first define the general notion of scattered sets.

2.9. Definition (PA). Let $A = \{a_0, \ldots, a_k\}$. Assume that F is a total function (definable in PA, e.g. $2_1^x = 2^x$, $2_2^x = 2^{2^x}$, 2_l^x, etc.). We say that

$$A \text{ is } F\text{-scattered iff } \forall i < k\; F(a_i) < a_{i+1}.$$

Σ_n-scattered sets can be in a sense considered as a special case of F-scattered sets. For $n \geqslant 1$ define $H^n(a) = b$ iff

$$\begin{aligned}&\forall \theta(\bar{x}, \bar{y}) \in \Pi_{n-1}\, [\ulcorner\exists \bar{y}\, \theta(\bar{x}, \bar{y})\urcorner < a \wedge \bar{a} < a \wedge \mathrm{Tr}_{\Sigma_n}(\exists \bar{y}\, \theta(\bar{a}, \bar{y})) \rightarrow \\ &\quad \exists \bar{b} < b\, \mathrm{Tr}_{\Pi_{n-1}}(\theta(\bar{a}, \bar{b}))] \wedge \forall z < b\, \exists \bar{a} < a\, \exists \theta < a [\ulcorner\exists \bar{y}\, \theta(\bar{x}, \bar{y})\urcorner < a \wedge \\ &\quad \exists \bar{b} < b\, \mathrm{Tr}_{\Pi_{n-1}}(\theta(\bar{a}, \bar{b})) \wedge \forall \bar{b} < b\, (\mathrm{Tr}_{\Pi_{n-1}}(\theta(\bar{a}, \bar{b})) \rightarrow \neg \bar{b} < z)]\end{aligned}$$

where $\bar{a} < a$ means that all the terms of the sequence $\bar{a}$ are less than a.

H^n is obviously of class Π_n in $\mathrm{I}\Sigma_n$ and is total in $\mathrm{I}\Sigma_n$ (because the strong Σ_n-collection principle is provable in $\mathrm{I}\Sigma_n$). Moreover, H^n dominates all total functions of class Σ_n; this is provable in $\mathrm{I}\Sigma_n$. For $n = 0$ we make the convention $H^0(x) = x + 1$.

It is obvious that every H^n-scattered set is Σ_n-scattered. One can also show that conversely, *every* Σ_n-scattered set A with a_0 large enough is H_n-scattered. We omit the proof vecause this is not essential for our purposes.

Now let $A^{\geqslant r}$ denote the set obtained from A by deleting the r final elements.

2.10. Proposition. *Assume that* $F(x, y) \in \Sigma_1$ *is a formula which defines a total function* $y = F(x)$ *in* PA *and let* $r = \mathrm{a.r.}(F)$. *Then the following implication is valid: if* $\mathrm{PA} \vdash_m \forall x\, \exists y\, F(x, y)$ *then*

$$\mathrm{I}\Delta_0 + \exp \vdash \forall A\ (\text{“}A \textit{ is an } (x+1)^2\textit{-scattered } \underline{m}\textit{-skeleton} \\ \rightarrow A^{\geqslant r} \textit{ is } F\textit{-scattered”}).$$

Proof. Let $F(x, y) := \exists \bar{z}\, F'(x, y, \bar{z})$, where $F' \in \Delta_0$. Assume that $\mathrm{PA} \vdash_m \forall x\, \exists y\, F(x, y)$. By Σ_1 completeness, $\mathrm{I}\Delta_0 + \exp$ proves $\mathrm{PA} \vdash_m \forall x\, \exists y\, \ulcorner F(x, y) \urcorner$ (we identify the Gödel number with a term number). By the Soundness Lemma, $A \models \forall x\, \exists y\, \exists \bar{z}\, \ulcorner F'(x, y, \bar{z}) \urcorner$ and by the definition of the relation $\models$ (compare also Proposition 1.9)

$$\forall x < a_i\, \exists y, \bar{z} < a_{i+1}\, A \models \ulcorner F'(x, y, \bar{z}) \urcorner \quad \text{for } i < k - r.$$

By absoluteness of Δ_0-formula we conclude the proof of the proposition. □

Since all the functions H_α: $\alpha < \varepsilon_0$ (independently of the underlying system of sequences) are provably total in PA we obtain the following corollary.

2.11. Corollary. *For every* $\alpha < \varepsilon_0$ *there exists a constant c such that*

$$\mathrm{I}\Delta_0 + \exp \vdash \forall A \text{ “}A \textit{ is a } \underline{c}\textit{-skeleton} \rightarrow A^{\geqslant c} \textit{ is } H_\alpha\textit{-scattered”}.$$

2.12. Note. From now on throughout the paper we use term numbers constructed as follows:

$$\underline{0} = 0, \qquad \underline{1} = 1, \qquad \underline{2} = 1 + 1, \qquad \underline{3} = 1 + \underline{2}\,\underline{1}$$

$$\underline{2a+1} = 1 + \underline{2}\,\underline{a}, \qquad \underline{2a} = 0 + \underline{2}\,\underline{a} \quad \text{for } a > 1.$$

Moreover we assume that the alphabet of L_{PA} is a finite set of symbols, e.g. k symbols. We encode formulas by k-ary expansions. This guarantees the following three properties:

(1) $\mathrm{I}\Sigma_1 \vdash \forall \theta\, (\theta \in L_{\mathrm{PA}} \rightarrow \theta < k^{\mathrm{lh}(\theta)})$.

(2) There exists a constant c such that $I\Sigma_1 \vdash \forall a \geq 2\,(\underline{a} \leq a^c)$.
(3) For every $\varphi(x, R) \in L_{PA}(R)$ there exists a constant c such that

$$I\Delta_0 + \exp \vdash \forall \theta \in L_{PA}\, \forall a \geq 2\, (\ulcorner \varphi \urcorner(\underline{a}, \theta) \leq \max(a, \theta)^c).$$

For further considerations, in the present chapter and in Chapters II.4–V, it will be useful to enrich the language of informal formulations of proofs and theorems with the notion of 'concrete numbers'. This notion concerns the numbers which are defined 'outside' the theory by term numbers. We use this notion only when it is clear in which way we can translate a proof possessing such a phrase into a metaproof with the corresponding phrase 'term numbers'; and the same for properties and theorems. For example the property (3) can be written as follows.

$I\Delta_0 + \exp$ proves that every expression built of $\underline{a}$ and θ and symbols of the alphabet in a concrete number of steps is less than $\max(a, \theta)^c$ for some concrete number c.

We end this section with the following observation on how large a_0 is in comparison with a for a-skeletons.

2.13. Lemma. *For every function F, Σ_1 definable in* PA *there exists a term number c_0 such that*

$$I\Delta_0 + \exp \vdash \forall A\, (\text{“}A \textit{ is an a-skeleton”} \wedge a \geq c_0 \rightarrow F(a) \leq a_0).$$

Proof. We can assume without loss of generality that F is increasing in PA. Let m be a constant such that $\mathrm{PA} \vdash_m \forall x\, \exists y\, F(2^x) = y$. By Σ_1-completeness $I\Delta_0 + \exp$ proves $\mathrm{PA} \vdash_{\underline{m}} \forall x\, \exists y\, \ulcorner F(2^x) = y \urcorner$.

Now we work in $I\Delta_0 + \exp$. The concrete number c_0 will be chosen in the course of the proof. Take a number $a \geq 4$. Let $b = \ulcorner \log_2 a \urcorner + 1$ and assume that

$$\underline{b} = i_1 + (\underline{2}(i_2 + \underline{2}(\cdots))),$$

where $i_1, i_2, \ldots$ are 0 or 1. It is easy to construct a sequence $t_1, t_2, \ldots, t_l \leq b^{c_1}$ (for some concrete c_1) of terms such that $t_0 = i_1 + x$, $t_l = \underline{b}$ and that the successive terms are obtained by a substitution of simple terms: take

$$i_1 + x, \qquad i_1 + (z \cdot x), \qquad i_1 + ((z+1) \cdot x), \qquad i_1 + (\underline{2} \cdot x),$$
$$i_1 + \underline{2} \cdot (i_2 + x), \qquad i_1 + \underline{2} \cdot (i_2 + (z \cdot x)), \qquad \text{etc.}$$

Take a proof of the formula $\forall x\, \exists y\, \ulcorner F(2^x) = y \urcorner$ of width $\leq \underline{m}$. Denote $\ulcorner \exists y\, F(2^x) = y \urcorner$ by $\theta(x)$. Then p extended by the sequence

$$\forall x\, \theta(x), \qquad \forall x\, \theta(x) \rightarrow \theta(t_0(x)), \qquad \theta(t_0(x)), \qquad \forall x\, \theta(t_0(x)),$$
$$\forall x\, \theta(t_0(x)) \rightarrow \theta(t_1(x, z)), \qquad \theta(t_1(x, z)), \qquad \forall z\, \theta(t_1(x, z)), \ldots, \theta(t_l)$$

is a proof in Hilbert style as was adopted in Section II.1. The width of this proof is less than $\max\{\forall z\, \theta(t_{l-2}(x,z)) \rightarrow \theta(t_{l-1}(x)),\ \forall x\, \theta(t_{l-1}(x)) \rightarrow \theta(t_l), \underset{\sim}{m}\}$. Hence (cf. 2.12(3)) it is less than b^c for a concrete number c, i.e., $\mathrm{PA} \vdash_{b^c} \exists y\ \ulcorner F(2^b) = y \urcorner$. Finally we choose a concrete number c_0 such that $a \geq c_0$ implies $(1 + \log_2 a)^c \leq a$.

If A is an a-skeleton then by the Soundness Lemma the sentence $\exists y\ \ulcorner F(2^b) = y \urcorner$ is A-true. By 1.9(iii) it then follows that there exists $y < A_0$ such that $A \models \ulcorner F(2^b) = y \urcorner$, i.e. $F(2^b) = y$, which follows by an absoluteness argument. Thus $F(2^b) < a_0$, and this finishes the proof. □

2.14 Note (PA). Assume that $\theta(\bar{x}) \in \Delta_0(2^x)$ and let l be a number such that all terms in θ are less than 2^x_l. Assume that θ is naturally interpreted in L_{PA} as a Σ_1 formula. Then for every 2^x_l-scattered set A such that $|A| \geq \mathrm{a.r.}(\theta)$ we have the absoluteness $\forall \bar{x} < a_0\, [(A \models \theta(\bar{x})) \Leftrightarrow \mathrm{Tr}_{\Sigma_1}(\theta(\bar{x}))]$. It follows by 2.11 that there exists a constant c such that the truth for θ is absolute with respect to every c-skeleton A.

Moreover, observe that every Σ_1-formula $\theta(\bar{x})$ is upwards absolute, i.e., for every set A with $|A| \geq \mathrm{a.r.}(\theta) + i + 1$, we have

$$\forall \bar{x} < a_i\, [(A \models \theta(\bar{x})) \rightarrow \mathrm{Tr}_{\Sigma_1}(\theta(\bar{x}))].$$

II.3. α-large a-skeletons

This section constitutes the first step towards the construction of a-skeletons which are models for parts of $\mathrm{I}(\varepsilon_\alpha)$ and $\mathscr{R}_\alpha(\mathrm{PA})$.

Under appropriate assumptions we construct here ε_α-large a-skeletons (the notion of largeness is very similar to that of Ketonen–Solovay [12]). To make the full construction real, i.e. to prove the sentence $\forall a\, \exists A$ "A is an ε_α-large a-skeleton", our assumptions should be as strong as the sentence $\mathrm{CR}(\mathrm{I}(\varepsilon_\alpha); \Sigma_1)$ (hence, by 2.6, at least as strong as $\mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_1)$). This is a consequence of Lemma III.1.5, which will be proved in the next chapter and which says roughly that

$$\forall a\, \exists b_a \leq A\, \forall A\, [\text{"}A \text{ is an } \varepsilon_\alpha\text{-large } a\text{-skeleton"} \rightarrow A \vdash_{b_a} \mathrm{I}(\varepsilon_\alpha)]$$

and $\lim_{a \to \infty} b_a = \infty$. This lemma constitutes the second step of the construction of a-models for $\mathrm{I}(\varepsilon_\alpha)$.

Here we show that the sentence $\forall a\, \exists A$ "A is an ε_α-large a-skeleton" is in fact provable in $\mathrm{I}\Sigma_1 + \mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_1)$.

Before we formulate the main lemma we have to define precisely the notion of α-large sets. We use the notion very close to that of Solovay and Ketonen [12] but we introduce it in close connection with the Hardy hierarchy H_α: $\alpha < \lambda$.

Assume that λ defines a basic system of notations in $\mathrm{I}\Sigma_1$ (of class Σ_1) and let P define a system of sequences for λ in $\mathrm{I}\Sigma_1$.

We denote by S_A the function such that $\mathrm{dom}\, S_A = \{a_0, \ldots, a_{k-1}\}$ whenever $A = \{a_0, \ldots, a_k\}$ and $S_A(a_i) = a_{i+1}$ for $i < k$.

3.1. Definition ($\mathrm{I}\Sigma_1$). We say that a finite set A is α, *P-large* iff $(S_A)^P_\alpha(a_0)\downarrow$, where $(S_A)^P_\alpha$ denotes the αth iterate of S_A (see Definition I.2.3).

Shortly, we say A is α-large. Let $A = \{a_0, \ldots, a_k\}$. For example A is n-large iff $|A| > n$, A is ω-large iff $a_{\omega_{a_0}} \leqslant k$ but A is $\omega + 1$-large iff $a_{\omega_{a_1}} \leqslant k$.

3.2. Proposition ($\mathrm{I}\Sigma_1$). *Assume that f is a finite partial function with $\forall x \in \operatorname{dom} f\ x < f(x)$. Then the following implication is valid:*

$$f^P_\alpha(a)\downarrow \to \text{“}\{f^k(a): f^k(a) \leqslant f^P_\alpha(a)\} \textit{ is } \alpha\textit{-large} \wedge f\textit{-scattered”}.$$

Proof. Indeed, set $A = \{f^k(a): f^k(a) \leqslant f_\alpha(a)\}$. Then $(S_A)^k(a) = f^k(a)$ for all $k \leqslant l$, where $f^l(a) = f^P_\alpha(a)$. Since the witness sequence for $y = f_\alpha(a)$ is composed of the iterations $f^k(a)$, $k \leqslant l$, it follows that $f^P_\alpha(a) = (S_A)^P_\alpha(a)$, which finishes the proof (cf. the proof of I.2.4). □

One can easily show that the connection with α-large sets of Ketonen–Solovay [12] is as follows.

Note. (1) If A is α-K.S. large, then A is α-large.

(2) If A is $\alpha + 2$-large then A is α-K.S. large whenever P is a B^+-system of sequences.

We write $\alpha \Rightarrow_x \beta$ if there exist finite sequences $\alpha = \alpha^0, \alpha^1, \ldots, \alpha^l = \beta$ and $x_0, x_1, \ldots, x_l \leqslant x$ such that $(\alpha^i)^P_{x_i} = \alpha^{i+1}$ for $0 \leqslant i < l$. Obviously $\alpha \rightarrow_x \beta$ implies that $\alpha \Rightarrow_x \beta$. An immediate consequence of Lemma I.2.7(i) (on the local monotonicity of the hierarchy f_α: $\alpha < \lambda$) is the following proposition.

3.3. Proposition ($\mathrm{I}\Sigma_1$). *Assume that P defines in $\mathrm{I}\Sigma_1$ a B^+-system of sequences for λ. Then the following implication is valid:*

$$\text{“}A \textit{ is } \alpha\textit{-large”} \wedge (\alpha \Rightarrow_{a_0} \beta) \to \text{“}A \textit{ is } \beta\textit{-large”}.$$

Now to formulate the main lemma assume that an ε-system of notation λ in $\mathrm{I}\Sigma_1$ is fixed, and analogously a B^+-ε-system of sequences P for λ in $\mathrm{I}\Sigma_1$ is fixed. This is our standing assumption in the remainder of this section.

3.4. Lemma ($\mathrm{I}\Sigma_1$). (i) *If A is ε_0-large, $a_0 \geqslant 3$ then there exists an a_0-skeleton $B \subseteq A$.*

(ii) *If A is $\omega_m^{\varepsilon_\beta+1}$-large, $a_0 \geqslant m \geqslant 3$ then there exists an $\varepsilon_\beta + 1$-large m-skeleton $B \subseteq A$.*

(iii) *If A is $\varepsilon_{\beta+1}$-large then there exists an $\varepsilon_\beta + 1$-large a_0-skeleton $B \subseteq A$.*

The proof of the lemma takes up the rest of this section; now we consider some corollaries.

3.5. Note. Basing on 3.4 one can prove the simplest case of the Kreisel–Levy Theorem, i.e. the statement $\mathrm{I}(\varepsilon_0) \equiv \mathcal{R}_0(\mathrm{PA})$, without any further preparations, where $\mathcal{R}_0(\mathrm{PA}) \equiv \mathrm{PA} + \{\mathrm{R}(\mathrm{PA}; \Sigma_n)\colon n \in \omega\}$. The combinatorial-logical proof of the full theorem will be given in Chapter IV. $\mathrm{I}(\varepsilon_0) \subseteq \mathcal{R}_0(\mathrm{PA})$ is easy to check (see Gentzen [7] and it is enough to show $\mathrm{I}(\varepsilon_0) \vdash \mathrm{R}(\mathrm{PA}; \Sigma_n)$ for all n. Hence by 2.6 it is enough to show $\mathrm{I}(\varepsilon_0) \vdash \mathrm{CR}(\mathrm{PA}; \Sigma_n)$ for all n.

Fix n and work in $\mathrm{I}(\varepsilon_0)$. By 2.7 we have to prove $\forall a\, \exists A$ "A is a Σ_{n-1}-scattered a-skeleton". Let H^{n-1} be the function from the previous section which was introduced to characterize Σ_{n-1}-scattered sets. Arguing inductively we prove $\forall \beta \leqslant \varepsilon_0$ "$(H^{n-1})_\beta$ is total". By 3.2, the set $B = \{(H^{n-1})^k(a)\colon (H^{n-1})^k(a) \leqslant (H^{n-1})_{\varepsilon_0}(a)\}$ is ε_0-large and H^{n-1}-scattered. By 3.4(i) we infer that there exists an a-skeleton $A \subseteq B$, which finishes the proof because A is also obviously Σ_{n-1}-scattered. □

At the beginning of this section we discussed the simplest case (for $n = 1$) of the next corollary. We observe at once that arguments presented in the discussion can as well be used to prove in $\mathrm{I}\Sigma_1$ the opposite implication of 3.6.

3.6. Corollary. *Let* $n \geqslant 1$. *Then* $\mathrm{I}\Sigma_1$ *proves that* $\mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_n) \rightarrow \forall a\, \exists A$ "*A is ε_α-large $\wedge$ A is Σ_{n-1}-scattered a-skeleton*".

Proof. Fix n and work in $\mathrm{I}\Sigma_1 + \mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_n)$. Then fix a. Since $\mathrm{I}(\varepsilon_\alpha) \equiv \mathrm{I}(<\varepsilon_{\alpha+1})$ (see Gentzen [7]) we have

$$\mathrm{I}(\varepsilon_\alpha) \vdash \text{“}(H^{n-1})_{\omega_a^{\varepsilon_\alpha+1}} \textit{ is total”}.$$

By 3.4(ii) we infer that (cf. the proof in the above note)

$$\mathrm{I}(\varepsilon_\alpha) \vdash \exists A\,[\text{“}A \text{ is an } \varepsilon_\alpha + 1\text{-large and } \Sigma_{n-1}\text{-scattered } \underline{a}\text{-skeleton}].$$

The composite notion "A is $\varepsilon_{\underline{\alpha}+1}$-large" is of class Σ_1 because so are the "components" "A is β-large" and "$\beta = \varepsilon_{\underline{\alpha}+1}$". Next, the notion "$A$ is Σ_{n-1}-scattered" is of class Π_{n-1}. Thus the formula on the right-hand side of the sign $\vdash$ is of class Σ_n. Hence the principle $\mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_n)$ can be used and we obtain the desired conclusion. □

3.7. Corollary. $\mathrm{I}(\varepsilon_{\alpha+1})$ *proves that* $\forall a\, \exists A$ "*A is an ε_α-large and Σ_{n-1}-scattered a-skeleton*".

This is an obvious corollary from 3.6, because Kreisel–Levy [14] proved that $\mathrm{I}(\varepsilon_{\alpha+1}) \vdash \mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_n)$. Their proof has a proof-theoretic character. An independent combinatorial-logical proof of 3.7 (as well as of the Kreisel–Levy Theorem, see Section IV.1) is also possible.

Indeed, first we show

$$\mathrm{I}(\varepsilon_{\alpha+1}) \vdash \forall a \text{ “}(H^{n-1})_{\omega_a^{\varepsilon_\alpha+1}} \text{ is total”}.$$

Then arguing in the same way as in 3.5 and 3.6 we get the desired conclusion. □

The proof of 3.4 will be preceded by a series of preparatory lemmas.

3.8. Lemma ($\mathrm{I}\Sigma_1$). *For every α and for every n the following implication is valid:* $\varepsilon_\delta < \alpha < \varepsilon_{\delta+1} \rightarrow \varepsilon_\delta \leqslant \alpha_n$.

This lemma is almost obvious for B^+-ε-systems of sequences but we show it using only properties (i)–(iii) of Definition I.1.8.

Proof. We work in $\mathrm{I}\Sigma_1$. Assume to the contrary that for some α and n, $\alpha_n < \varepsilon_\delta < \alpha < \varepsilon_{\delta+1}$. Take the α satisfying the above inequalities and having the smallest code $\ulcorner\alpha\urcorner$.

By Definition I.1.4 of ε-systems of notations we have the following possibilities.

Case (a): $\alpha = \beta \dotplus \gamma$, *where* $\gamma > 0$. The property I.1.8(i) reads $(\beta \dotplus \gamma)_n = \beta \dotplus \gamma_n$ for $\gamma > 0$. Hence $\alpha_n = \beta \dotplus \gamma_n > \varepsilon_\delta$, a contradiction.

Case (b): $\alpha = \omega^\beta$, *where* $\beta = \gamma + 1$. Then $\gamma \geqslant \varepsilon_\delta$. The property I.1.8(ii) reads $(\omega^{\gamma+1})_n = \omega^\gamma(n+1)+1$ and obviously $\alpha_n = \omega^\gamma(n+1)+1 > \varepsilon_\delta$, a contradiction.

Case (c): $\alpha = \omega^\beta$, *where* $\beta \in Lim$ *and* $\ulcorner\beta\urcorner < \ulcorner\alpha\urcorner$. The property I.1.8(iii) reads $(\omega^\beta)_n = \omega^{\beta_n}$. Hence $\beta_n < \varepsilon_\delta < \beta$ and we obtain a contradiction. $\square$

3.9. Lemma ($\mathrm{I}\Sigma_1$). *If A is $\omega^3(\alpha+1)$-large and $a_0 \geqslant 3$ then there exists a 2_3^x-scattered and α-large set $B \subseteq A$.*

Proof. Now we work informally in the system $\mathrm{I}\Sigma_1$. Let A be an arbitrary set with $a_0 \geqslant 3$. We set $f = S_A$ and define $g(x) = f_{\omega^3}(x)$. It is easy to check that $g(x) \geqslant 2_3^x$ for $x \geqslant 3$, if $g(x)\downarrow$.

We will show by induction on β (which in fact can be treated as finite induction, by Lemma I.2.6 on finitization) the following claim.

(1) **Claim.** *For all β and for all $x \in A$,*

$$g_\beta(x) \leqslant f_{(\omega^3+1)\beta}(x).$$

Hence if A is $\omega^3(\alpha+1)$-large then $g_\alpha(a_0)\downarrow$. It follows that $B = \{g^k(a_0)\colon g^k(a_0)\downarrow \wedge k \text{ is a natural number}\}$ is α-large and 2_3^x-scattered, which is the desired conclusion of our lemma.

Now we prove the claim. If $\beta = 0$ then the claim is obvious. Now we consider the inductive step $\beta \rightarrow \beta+1$. Let $x \in A$. Then

$$f_{\omega^3(\beta+2)}(x) \simeq f_{\omega^3(\beta+1)}(f_{\omega^3}(x)) \simeq f_{\omega^3(\beta+1)}(g(x)) \geqslant g_\beta(g(x)) \simeq g_{\beta+1}(x),$$

which finishes the proof of this step.

Assume now that (1) is true for ordinals $<\beta$, where $\beta \in Lim$. Let $x \in A$ and $f_{\omega^3(\beta+1)}(x)\downarrow$. Then

$$(2) \qquad f_{\omega^3(\beta+1)}(x) \simeq f_{\omega^3\beta}(g(x)).$$

Denote $y = g(x)$. Now we show that

(3) $\omega^3\beta \Rightarrow_y \omega^3 \cdot \beta_x$.

We can represent β in the form $\gamma \dot{+} \omega^\delta$. For $\delta = 0$ the point (3) is obvious. Hence we have to consider four cases.

Case (i): $\delta = m + 1 \leq \omega$. Then $\omega^3(\gamma \dot{+} \omega^{m+1}) \Rightarrow_y \omega^3\gamma \dot{+} \omega^{m+3}(y+1) + 1 \Rightarrow_y \omega^3\gamma \dot{+} \omega^{m+3}(x+1) \dot{+} \omega^3 = \omega^3(\gamma \dot{+} \omega^{m+1})_x$, because $y \geq x + 1$.

Case (ii): $\delta = \omega$. Then $\omega^3(\gamma \dot{+} \omega^\omega) \Rightarrow_y \omega^3\gamma \dot{+} \omega^{y+2} \Rightarrow_0 \omega^3\gamma \dot{+} \omega^{x+5} = \omega^3(\gamma \dot{+} (\omega^\omega)_x) = \omega^3(\gamma \dot{+} \omega^\omega)_x$, because $y \geq x + 3$.

Case (iii): *$\delta > \omega$ and δ is limit.* Then $\omega^3(\gamma \dot{+} \omega^\delta) \Rightarrow_y \omega^3\gamma \dot{+} \omega^{\delta_y} \Rightarrow_0 \omega^3\gamma \dot{+} \omega^{\delta_x} = \omega^3\gamma \dot{+} \omega^3\omega^{\delta_x} = \omega^3(\gamma \dot{+} \omega^\delta)_x$, because $\delta_x \geq \omega$ and $\delta_y \Rightarrow_0 \delta_x$ which follows by I.2.6(iii) and the fact that P is a B^+-system.

Case (iv): *$\delta > \omega$ and $\delta = \eta + 1$.* Then $\omega^3(\gamma \dot{+} \omega^\delta) \Rightarrow_y \omega^3\gamma \dot{+} \omega^\eta(y+1) + 1 \Rightarrow_0 \omega^3\gamma \dot{+} \omega^\eta(x+1) \dot{+} \omega^3 = \omega^3(\gamma \dot{+} \omega^{\eta+1})_x$, because $\omega^\eta \Rightarrow_0 \omega^3$.

By (3) and I.2.7(i) it follows that

$$f_{\omega^3\beta}(y) \geq f_{\omega^3\beta_x}(y) = f_{\omega^3(\beta_x+1)}(x).$$

By the inductive assumption, this is $\geq g_{\beta_x}(x) = g_\beta(x)$, which by (2) finishes the proof of the limit step. □

Assume that f is a finite partial function from a subset of $\mathbb{N}$ to $\mathbb{N}$. We say that (A, b) approximates f iff

$$\max A \leq b \wedge \forall a \in A - \{\max A\}\ \forall x < a - 2\,[f(x) < a^+ \vee b \leq f(x)],$$

where $a^+ = \min\{x \in A : a < x\}$.

A approximates f iff $(A, \max A)$ approximates f.

3.10. Lemma ($I\Sigma_1$). *For all $\alpha, \beta < \lambda$ such that $\alpha \dot{+} \beta\downarrow$, for each $\omega^{\alpha\dot{+}\beta}$-large set A with $\alpha \neq 0$, $a_0 \neq 0$ and for each f there exists an β-large set $B \subseteq A$ with $b_0 = a_0$ which approximates f and there exists a $b \in A$ such that* $[\max B, b] \cap A$ *is ω^α-large.*

The special case of this lemma for $\lambda = \varepsilon_0$ was proved in [23]. For the proof see also [10]. It was observed in [11] that if we strengthen the assumption "A is $\omega^{\alpha\dot{+}\beta}$-large" to "$\alpha \neq 0 \wedge A$ is $\omega^{\alpha\dot{+}\beta}$-large" then the proof from [10] (by induction on β) still works verbatim if we replace ε_0 by λ', provided the system of sequences for λ' satisfies the following axioms:

$$(\omega^{\alpha+1})_n = \omega^\alpha \cdot n, \qquad (\omega^\alpha)_n = \omega^{\alpha_n} \quad \text{for all } \alpha \in Lim\backslash eps,$$

$$(\alpha + \beta)_n = \alpha + \beta_n \quad \text{for } \alpha, \beta \text{ such that } \alpha \dot{+} \beta\downarrow.$$

Our B^+-ε-system λ satisfies the axiom $PA \vdash (\omega^{\alpha+1})_n = \omega^\alpha(n+1) + 1$, hence the above observation does not directly apply.

Before we turn to the proof of this lemma we observe that after putting in it ω^α in place of α and of β we obtain the following lemma more appropriate for our purposes.

3.11. Lemma ($I\Sigma_1$). *For each $\alpha < \lambda$ such that $\omega^{\omega^{\omega^\alpha \cdot 2}} < \lambda$ and for each $\omega^{\omega^{\omega^\alpha \cdot 2}}$-large set A with $a_0 > 0$ and for each function f there exists an ω^α-large set $B \subseteq A \setminus \{\max A\}$ which approximates f.*

The $I\Sigma_1$-proof of 3.10 is obtained by a finitization of the proof from [10].

Proof of Lemma 3.10. Our lemma has the following form:

$$\forall \alpha\, \forall \beta\, \forall A\; T(\alpha, \beta, A).$$

Now we work in $I\Sigma_1$. The above statement is in an obvious way equivalent to the statement

$$\forall a \geq 1\, \forall \alpha\, \forall \beta\, \forall A \subseteq [1, a]\; T(\alpha, \beta, A).$$

For $a \geq 1$, let $X_a = \{\gamma \colon \exists A \subseteq [1, a]\; A \text{ is } \gamma\text{-large}\}$. By I.2.6, the set X_a is finite. Let $Y_a = \{\beta \colon \exists \alpha \neq 0\; \omega^{\alpha \dot{+} \beta} \in X_a\}$. Since $\ulcorner \beta \urcorner < \ulcorner \alpha \dot{+} \beta \urcorner < \omega^{\ulcorner \alpha \dot{+} \beta \urcorner}$ for α and $\beta \neq 0$ such that $\alpha \dot{+} \beta \downarrow$, we infer that Y_a is also finite. Using Y_a we can equivalently formulate our statement as follows:

$$\forall a \geq 1\, \forall \beta \in Y_a\, \forall \alpha\, \forall A \subseteq [1, a]\; T(\alpha, \beta, A).$$

Fix $a \geq 1$. To show the above statement we use induction on $\beta \in Y_a$. The case $\beta = 0$ is obvious.

Case 1: $0 \neq \beta \in Lim \wedge \beta \in Y_a$. Fix $\alpha \neq 0$. Since $\beta \in Lim$, the following implication is true: $T(\alpha, \beta_{a_0}, A) \to T(\alpha, \beta, A)$. If $\beta_{a_0} \notin Y_a$ then A is not $\omega^{\alpha \dot{+} \beta_{a_0}}$-large, hence A is not $\omega^{\alpha \dot{+} \beta}$-large and $T(\alpha, \beta, A)$ is obviously true. In the other case the conclusion for β follows from the inductive hypothesis.

Case 2: $\beta = \gamma + 1 \wedge \beta \in Y_a$. Let $A \subseteq [1, a]$ be $\omega^{\alpha \dot{+} \beta}$ for a certain α. Let f be a function. Hence A is also $\omega^{\alpha \dot{+} \gamma} \cdot a_0$-large. Let $a_k = (S_A)_{\omega^{\alpha \dot{+} \gamma} \cdot k}(a_0)$ for $k = 0, \ldots, a_0$. One can easily check, using suitable finitization of transfinite induction that $a_{k+1} = (S_A)_{\omega^{\alpha \dot{+} \gamma}}(a_k)$ for $k < a_0$. (Generally it is true that if $\alpha^1 + \alpha^2 \downarrow$, then $g_{\alpha^1 \dot{+} \alpha^2} = g_{\alpha^1} \circ g_{\alpha^2}$; cf. [32] and [10].) Obviously $a_0 < \cdots < a_{a_0}$. By the Pigeon-Hole Principle there exists a j_0 with $1 \leq j_0 \leq a_0 - 1$ and $[a_{j_0}, a_{j_0+1}) \cap f * [0, a_0 - 2) = \emptyset$ because there are at most $a_0 - 2$ images $f(x)$ of $x < a_0 - 2$ and there are $a_0 - 1$ corresponding intervals.

We let $A' = A \cap [a_{j_0}, a_{j_0+1}]$. Then A' is $\omega^{\alpha \dot{+} \beta}$-large, since $(S_A)_{\omega^{\alpha \dot{+} \gamma}}(a_{j_0}) = a_{j_0+1}$. Moreover $\gamma \in Y_a$. By the inductive assimption there exists an γ-large $B' \subseteq A'$ with $b'_0 = a_{j_0}$ which approximates f and there exists a $b \in A'$ such that $[\max B', b] \cap A$ is ω^α-large. Thus the pair $B = B' \cup \{a_0\}$ and b has the desired properties. □

The proof of Lemma 3.4 is preceded by one more helpful lemma, which provides a sort of reduction of the construction of sets of diagonally indiscernible elements to the construction of approximate sets. We denote temporarily by $a^{(n)}$ the last but nth element of A, $a^{(0)} = \max A$.

3.12. Lemma ($\mathrm{I}\Delta_0 + \exp$). *Assume that we are given a sequence $\eta_i(\bar{x}_i, \bar{y}_i)$: $i < k$ υ_s formulas such that* a.r.$(\eta_i) = n$ *for $i < k$ and* a.r.$(\exists \bar{y}_i\, \eta_i(\bar{x}_i, \bar{y})) = n + 1$ *for $i < k$. Moreover, assume that* $\mathrm{lh}(\bar{x}_i) \leq r$ *for $i < k$. Then to every $(k, x^r) + 2$-scattered set A we can associate a partial function f such that the following holds: if A is a set of diagonally indiscernible elements for all η_i^*: $i < k$, then for each B' included in A minus its last n elements which approximates f, every set $B \subseteq B'$ such that $\forall c, d \in B\, [c < d \rightarrow \exists b \in B'\, (a < b < d)]$ is a set of diagonally indiscernible elements for all $(\exists \bar{y}_i\, \eta_i(\bar{x}_i, \bar{y}_i))^*$: $i < k$.*

Proof. We prove the lemma informally in $\mathrm{I}\Delta_0 + \exp$.

Let $\langle \bar{x} \rangle$ be a function coding every sequence of length less than r by a number less that $(a + 1)^r$, where a is equal to the maximum of terms of this sequence. Let $d(x, i)$ denote the decoding function. Let $(x)_0$, $(x)_1$ denote the decoding functions of (x, y). We define $l(i) = \mathrm{lh}(\bar{x}_i)$ for $i = 0, \ldots, k - 1$. The function f is defined as follows:

$$f(x) = \min_{z \leq a^{(n)}} [\exists \bar{y} \leq z\, A \models \eta_{(x)_0}(d((x_1)_1, 0), \ldots, d((x)_1, l((x)_0)), \bar{y})].$$

Let B' included in A minus its last n elements be a set which approximates f, $B' = \{b_0, \ldots, b_l\}_<$. From the inequalities $i < k$, $\bar{x}_i < b_m$ it follows that $(i, \langle \bar{x}_i \rangle) < b_{m+1} - 2$, since A is $(k, x^r) + 2$-scattered. Therefore for every $i < k$ and for each $m \leq l - 2$

$$\forall \bar{x}_i < b_m\, [\exists \bar{y}_i < b_{m+2}\, A \models \eta_i(\bar{x}_i, \bar{y}_i) \Leftrightarrow \exists \bar{y}_i < b_l\, A \models \eta_i(\bar{x}_i, \bar{y}_i)].$$

Let $B \subseteq B'$ be such that $\forall c, d \in B\, [c < d \rightarrow \exists b \in B'\, (a < b < d)]$. In particular, we see that for all $b, b', b'' \in B$ with $b < b', b''$

$$(*) \qquad \forall \bar{x}_i < b\, [\exists \bar{y}_i < b'\, A \models \eta_i \Leftrightarrow \exists \bar{y} < b''\, A \models \eta_i].$$

We claim that B is a set of diagonally indiscernible elements for $(\exists \bar{y}_i\ \ \eta_i)^*$, where $i = 0, \ldots, k - 1$. To show this for fixed i let $b < D_1$, $D_2 \subseteq B$, $b \in B$ and $|D_1| = |D_2| = n + 1$. Denote by b', b'' the first elements of the sets D_1, D_2 respectively. In particular, $(*)$ is satisfied by b, b', b''. Since A is a set of diagonally indiscernible elements for η_i^* and a.r.$(\eta_i) = n$, we infer that the symbol A in $(*)$ to the left of $\Leftrightarrow$ can be replaced by $D_1 \setminus \{b'\}$, and the other A by $D_2 \setminus \{b''\}$. Hence finally

$$\forall \bar{x}_i < b\, [D_1 \models \exists \bar{y}_i\, \eta_i \Leftrightarrow D_2 \models \exists \bar{y}_i\, \eta_i]$$

and this shows that B is a set of diagonally indiscernible elements for $(\exists \bar{y}_i\, \eta_i)^*$, where $i < k$. $\square$

Proof of Lemma 3.4. We only prove (ii), the proof of (iii) is then immediate. The proof of (i) is similar.

Assume that A is $\omega_m^{\varepsilon_\beta+1}$-large, $a_0 \geq m \geq 3$. By 1.2.6(iii) we see that $\omega_m^{\varepsilon_\beta+1} \to_0 0$. By 3.8, $\omega_m^{\varepsilon_\beta+1} \to_0 \varepsilon_\beta \to_0 0$. We have $(\omega_m^{\varepsilon_\beta+1})_1 = \omega_{m-1}^{\varepsilon_\beta \cdot 2+1}$ and since $\varepsilon_\beta \to_2 3$ we infer that $\omega_m^{\varepsilon_\beta+1} \to_2 \omega_{m-1}^{\varepsilon_\beta+3}$. Set $\alpha = \omega^{\varepsilon_\beta+2}$. Now, $(\omega_{m-1}^{\varepsilon_\beta+3})_0 = \omega_{m-2}^{\alpha+1}$, $(\omega_{m-2}^{\alpha+1})_0 = \omega_{m-3}^{\omega^\alpha+1}$, etc. Summing up we have

$$\omega_m^{\varepsilon_\beta+1} \to_2 \omega_{m-1}^{\varepsilon_\beta+3} \to_0 \omega^{\omega_{m-3}^{\alpha}+1}.$$

Moreover, $\omega^{\omega_{m-3}^{\alpha}+1} \Rightarrow_1 \omega_{m-2}^{\alpha} \cdot 2$ and $\omega_{m-2}^{\alpha} \Rightarrow_2 \omega_{m-2}^{3} \Rightarrow_2 \omega^3$. Hence finally

(1) $$\omega_m^{\varepsilon_\beta+1} \Rightarrow_2 \omega_{m-1}^{\varepsilon_\beta+2} \dotplus \omega^3.$$

It follows by 3.3 that A is $\omega^3(\omega_{m-1}^{\varepsilon_\beta+2}+1)$-large. By Lemma 3.9 there exists an $\omega_{m-1}^{\varepsilon_\beta+2}$-large and 2_3^x-scattered set $C \subseteq A$. Hence without loss of generality we can assume that A itself is such a set.

Now we define by induction a helpful sequence of mappings $\gamma_k(\alpha)$:

$$\gamma_0(\alpha) = \alpha, \qquad \gamma_{k+1}(\alpha) = \omega^{\gamma_k(\alpha)\cdot 2}, \quad \alpha \text{ is here a variable.}$$

(2) **Claim.** *A is $\gamma_{m-2}(\omega^{\varepsilon_\beta+1})$-large.*

Indeed, set $\alpha = \omega^{\varepsilon_\beta+1}$. We prove by induction on $k \geq 1$ that $\omega_{k+1}^{\varepsilon_\beta+2} \Rightarrow_1 \gamma_k(\alpha)\cdot 2+1$. For $k=1$ this is obvious.

The inductive step is the following:

$$\omega_{k+1}^{\varepsilon_\beta+2} \Rightarrow_1 \omega^{\gamma_{k-1}(\alpha)\cdot 2+1} \Rightarrow_1 \omega^{\gamma_{k-1}(\alpha)\cdot 2}\cdot 2+1 = \gamma_k(\alpha)\cdot 2+1,$$

which finishes the proof of (2).

Now let us denote by Φ the set of all η which are generalizations of some $\theta < m$. Since the number of such η for every fixed θ is less than $2^{\mathrm{lh}(\theta)} \leq \theta$, it follows that $|\Phi| \leq m^2$.

To finish the proof of the Lemma we show the following claim.

(3) **Claim.** *For every n, every k and every $\gamma_{2n+k}(\omega^{\varepsilon_\beta+1})$-large and 2_3^x-scattered set A with $a_0 \geq m \geq 3$ there exists a $\gamma_k(\omega^{\varepsilon_\beta+1})$-large set B included in A minus its last n elements which is a set of diagonally indiscernible elements for all $\theta \in \Phi$ such that* a.r.$(\theta) \leq n$.

Indeed, this finishes the proof: we set $n = [(m-2)/2]$. Observe that for every $\eta \in \Phi$, a.r.$(\eta) \leq n$ (every natural coding of formulas guarantees this property). If we put $k=0$, we obtain a $B \subseteq A$ which is $\omega^{\varepsilon_\beta+1}$-large and is a set of diagonally indiscernible elements for θ^* for all $\theta \in \Phi$ such that a.r.$(\theta) \leq n$. Since $\omega^{\varepsilon_\beta+1} \to_2 \varepsilon_\beta+2$, B is also $\varepsilon_\beta+2$-large. Thus $B \setminus \{b_0\}$ is $\varepsilon_\beta+1$-large and it is also a set of indiscernible elements for θ^*, for all θ as above which are sentences. Hence finally $B \setminus \{b_0\}$ is an $\varepsilon_\beta+1$-large m-skeleton.

We show Claim (3) by induction on n. The case $n=0$ is obvious, since each set is a set of diagonally indiscernible elements for all θ^* such that a.r.$(\theta)=0$. Set $\alpha = \omega^{\varepsilon_\beta+1}$. Assume that the claim holds for n and assume that A is $\gamma_{2(n+1)+k}(\alpha)$-large and 2_3^x-scattered. By the inductive hypothesis there exists a $\gamma_{k+2}(\alpha)$-large set C included in A minus its last n elements which is a set of diagonally indiscernible elements for θ^* for all $\theta \in \Phi$ such that a.r.$(\theta) \leq n$.

Let $\theta_1, \theta_2, \ldots, \theta_l$, where $l \leqslant m^2$, be an enumeration of all $\theta \in \Phi$ with a.r.$(\theta) = n+1$ which are of the form $\exists \bar{y}\, \eta(\bar{x}, \bar{y})$, where a.r.$(\eta) = n$. Let $\theta_i := \exists \bar{y}_i\, \eta_i(\bar{x}_i, \bar{y}_i)$ and a.r.$(\eta_i) = n$ for all $i < l$. Certainly every set B which is a set of diagonally indiscernible elements for all θ_i^*: $i < l$ is diagonally indiscernible for all $\theta \in \Phi$ such that a.r.$(\theta) = n+1$.

Since A is 2_3^x-scattered and $a_0 \geqslant m \geqslant 3$, the set C is $(m, x^l) + 2$-scattered and Lemma 3.12 is applicable to C and to the sequence $\eta_1, \ldots, \eta_l$.

Let f be a function from 3.12 chosen for the sequence $\eta_i(\bar{x}_i, \bar{y}_i)$: $i < l$. Since C is $\gamma_{k+2}(\alpha)$-large, by Lemma 3.11 there exists a $\gamma_{k+1}(\alpha)$-large set $B' \subseteq C \setminus \{\max C\}$ which approximates f. Applying once more Lemma 3.11, we find a $\gamma_k(\alpha)$-large set $B \subseteq B'$ such that $\forall c, d \in B\, [c < D \rightarrow \exists b \in B'\, (c < b < d)]$. Namely, B approximates g such that $g(b-3) = b^+$ for $b \in B'$, where b^+ is the successor in B'. By Lemma 3.12, B is a set of diagonally indiscernible elements for all $\theta_i^*: i < l$. Obviously b is included in A minus its last $n+1$ elements, and this finishes the proof of Claim (3) and the proof of the lemma. □

3.13. Note. (a) Denote $\lg^3[B] = \{\max_y 2_3^y \leqslant b : b \in B\}$. We say that B is an (n, m)-skeleton if B is a set of diagonally indiscernible (or indiscernible) elements for θ^* for all $\theta < m$ such that a.r.$(\theta) \leqslant n$. Lemma 3.4(ii) can be refined as follows:

$$\exists \text{ concrete } c\, \forall A\, [A \text{ is } \omega_n^{\varepsilon_\beta \cdot 2+2}\text{-large} \wedge a_0 \geqslant m \geqslant c$$
$$\rightarrow \exists B \subseteq A\, (\text{“}B \text{ is } \varepsilon_\beta\text{-large} \wedge \lg^3[B] \text{ is an } (n, m)\text{-skeleton”})].$$

Reasoning similarly to the proof of (1), we show that we can assume that A is $\omega_n^{\varepsilon_\beta \cdot 2+1}$-large and 2_3^x-scattered. Then what remains to be done is a reduction of the number of references to 3.11 to one in every inductive step. This is possible if we refine 3.12 according to the ideas used in the proof of II.4.2.

(b) For systems of sequences satisfying

$$\forall \alpha < \lambda\, \forall n\, (\varepsilon_\alpha)_{n+1} \rightarrow_n \omega^{(\varepsilon_\alpha)_n}$$

a better result is possible. Namely we can weaken the assumption “A is $\omega_{n-1}^{\varepsilon_\beta(a_0+2)}$-large” to the form “$A$ is $\omega_{n-1}^{\varepsilon_\beta(a_0+2)}$-large”.

II.4. Partitions properties and a-skeletons

In this section we strengthen a little the construction of diagonally indiscernible elements of Paris–Harrington [21].

Let $[X]^e$ denote the family of all increasing sequences of the set X, having length e. For every $F : [X]^e \rightarrow r$ we write $H \in \mathrm{Hom}(F)$ instead of $H \subseteq X \wedge F \restriction [H]^e \equiv \mathrm{Const}$. Hence the Ramsey partition property $X \rightarrow_* (k)_r^e$ can be written as follows:

$$\forall (F : [X]^e \rightarrow r)\, \exists H \in \mathrm{Hom}(F)\, (|H| \geqslant \max(h_0, e+1, k)).$$

By careful inspection of the proof of Theorem (6) of [21] one can isolate the lemma formulated below which indicates how the existence of homogeneous sets 'forces' the existence of a-skeletons (in our terminology). We use the notation

$$\lg^i(x) = \max(\{y\colon 2_i^y \leqslant x\} \cup \{0\});$$

this is the natural extension of the ith iterate of $\log_2(x)$ (well defined only for the numbers 2_i^n: n natural) to all natural numbers. If $|H| \geqslant h_0$ we also say that H is relatively large. Define

$$\lg^i[H] = \{\lg^i(x)\colon x \in H\}.$$

4.1. Lemma (PA). *To every a and m we can associate a partition $P\colon [m]^e \to r$, where $e \geqslant a$ and r depend only on a, such that if H is homogeneous for P ($H \in \mathrm{Hom}(P)$) and $|H| \geqslant \max(h_0, e+1)$, then the set $\lg^4(H)$ minus its last e elements is an a-skeleton.*

Probably one can get a better result including concrete estimates of e and r. Here we present a more direct proof of 4.1, which gives the best estimate ($r = 2$, $e = a$) for suitable numbers a. In fact, we have the following lemma.

4.2. Lemma. *There exists a term number c such that $\mathrm{I}\Delta_0 + \exp$ proves the following.*

To every a and set A with $a_0 \geqslant c$, 2_4^a we can associate a partition P: $[A]^a \to 2$ such that if $H \in \mathrm{Hom}(P)$ and $|H| \geqslant h_0$, then the set $\lg^4[H \setminus \{h_0\}]$ minus its last a elements is an a-skeleton.

A similar refinement appears in [25]. This lemma will be used in Section IV.2 to study iterate partition properties. The proof of Lemma 4.2 is preceded by some preparatory lemmas. The first lemma says that some homogeneous sets (i.e., sets of indiscernible elements) are almost sets of diagonally indiscernible elements. Fix an n.

4.3. Lemma ($\mathrm{I}\Delta_0 + \exp$). *Let F be an arbitrary $\Delta_0(2^x)$-function bounded by 2_n^x. Let m be a number, A a set and R a relation included in $[0, \max A] \times [A]^m$.*

We define the relation $S_R \subseteq [A]^{2m+1}$ by

$$S_R(x, \bar{y}, \bar{z}) \Leftrightarrow \forall t < F(x)\, [R(t, \bar{y}) \Leftrightarrow R(t, \bar{z})].$$

Then it is true that for every set $H \subseteq A$ satisfying $m(2^{F(h_0)} + 1) < |H|$ which is homogeneous for the relation S_R the set $H' = H$ minus its last m elements has the following property:

$$\forall x \in H'\ \forall t < F(x)\ \text{“}H\text{-}[0, x] \textit{ is homogeneous for } R(t, \bar{y})\text{”}.$$

(*We will say that H' is $F(x)$-diagonally homogeneous for R.*)

The proof relies on the idea used by Paris and Harrington [21] (the proof of 2.10(iii)) and we omit it.

The next lemma says that some diagonally homogeneous sets are almost a-skeletons.

4.4. Lemma. *There exists a term number c_0 such that* $\mathrm{I}\Delta_0 + \exp$ *proves: for every set A and for every a with $a_0 \geq c_0$, 2_4^a there exists a relation $R \subseteq [0, \max A,] \times [A]^a$ such that for every set H which satisfies*

(1) $|H| \geq [h_0/2]$,

(2) *H is $\mathrm{lg}^2(x)$-diagonally homogeneous for $R(t, \bar{y})$,*

(3) *H is 2_4^x-scattered,*

the set $\mathrm{lg}^4[H \setminus \{h_0\}]$ is an a-skeleton.

Proof. The c_0 will be chosen in the course of the proof.

We work in $\mathrm{I}\Delta_0 + \exp$. For the construction of $R \subseteq [0, \max A] \times [A]^a$ (which forces the existence of a-skeletons) we enumerate all L_{PA}-formulas which are generalizations of formulas less than a: $\theta_0(\bar{x}^0), \ldots, \theta_l(\bar{x}^l)$. Set $m(i) = \mathrm{lh}(\bar{x}^i) - 1$ and $n(i) = \mathrm{lh}(\bar{y}^i) - 1$, where $\bar{y}^i$ is the sequence of new variables which appear in θ_i^* ($\theta_i^* = \theta_i^*(\bar{x}^i, \bar{y}^i)$). Every natural coding of formulas guarantees that $m(1), n(1), \ldots, m(l), n(l) < [a - 1/2] = m$ and that $l < a^2$.

Let $\langle \bar{x} \rangle$ be a function coding sequences of length less than a and with terms less than $\mathrm{lg}^4 b$ by numbers $< \sqrt{\mathrm{lg}^2 b}$ for $b \in A$. The existence of such a coding function is ensured if we choose c_0 in such a way that for all $b \geq c_0$, $(\mathrm{lg}^4 b)^{\mathrm{lg}^4 b} < \sqrt{\mathrm{lg}^2 b}$.

Let $d(y, i)$ denote the decoding function of $y = \langle \bar{x} \rangle$. Next, let (x, y) denote the usual polynomial pairing function and $(x)_0$, $(x)_1$ its decoding functions.

Now, we define the relation $R \subseteq [0, \max A] \times [A]^m$:

$$R(x, \bar{y}) \Leftrightarrow \mathrm{Tr}_0(\theta^*_{(x)_0}(d((x)_1, 0), \ldots, d((x)_1, m((x)_0)); \mathrm{lg}^4 y_0, \ldots, \mathrm{lg}^4 y_{n((x)_0)})).$$

It has the following property:

(∗) If H included in A is $\mathrm{lg}^2(x)$-diagonally homogeneous for R then $\forall b \in H$ $\forall \bar{x}^i < \mathrm{lg}^4 b$ "H-$[0, b]$ is homogeneous for $\mathrm{Tr}_0(\theta_i^*(\bar{x}^i; \mathrm{lg}^4 y_0^i, \ldots, \mathrm{lg}^4 y^i_{n(i)})$" for $i \leq l$.

Indeed, take H as above. Fix $b \in H$ and $i \leq l$. Finally, take $\bar{x}^i < \mathrm{lg}^4 b$. Hence $\langle \bar{x}^i \rangle < \sqrt{\mathrm{lg}^2 b}$ (because $\mathrm{lh}(\bar{x}^i) \leq a$) and therefore $x = (i, \langle \bar{x}^i \rangle) < \mathrm{lg}^2 b$, because $i \leq l \leq a^2 \leq (\mathrm{lg}^4 b)^2$.

By the choice of H and of x, H-$[0, b]$ is homogeneous for $R(x, \bar{y})$, but $R(x, \bar{y}) \equiv \mathrm{Tr}_0(\theta_i^*(\bar{x}^i, \mathrm{lg}^4 y_0, \ldots, \mathrm{lg}^4 y_{n(i)}))$ and (∗) follows.

This means that $\mathrm{lg}^4 H - [0, \mathrm{lg}^4 b]$ is homogeneous for $\mathrm{Tr}_0(\theta_i^*(\bar{x}^i, \bar{y}^i))$ for all $\bar{x}^i < \mathrm{lg}^4 b$. In particular $\mathrm{lg}^4[h \setminus \{h_0\}]$ is homogeneous for θ_i^* for all θ_i which are sentences. Therefore if H is 2_4^x-scattered and $|H| \geq [h_0/2]$, then by the definition of skeletons $\mathrm{lg}^4 [H \setminus \{h_0\}]$ is an a-skeleton, because

$$|\mathrm{lg}^4[H \setminus \{h_0\}]| = |H| - 1 \geq [h_0/2] - 1 \geq \mathrm{lg}^4 h_0 \geq a. \qquad \square$$

Now we can prove Lemma 4.2. Let c_0 be the constant from Lemma 4.4. The constant $c \geq c_0$ will be chosen in the course of the proof in $\mathrm{I}\Delta_0 + \exp$. Assume that the set A satisfies $a_0 \geq c, 2_4^a$. We put $m = [(a-1)/2]$. Let $R(x, \bar{y})$ be the relation from Lemma 4.4 included in $[0, \max A] \times [A]^m$ such that

(1) for every $\lg^2(x)$-diagonally homogeneous set $H \subseteq A$ satisfying $|H| \geq [h_0/2]$ the set $\lg^4[H \setminus \{h_0\}]$ is an a-skeleton, but under the assumption that H is 2_4^x-scattered.

We do not know whether H is 2_4^x-scattered but there exist four relations:

$$R_0(y_0, y_1) \equiv [1/2y_0] \leq y_1, \qquad R_1(y_0, y_1) \equiv 4 \cdot y_0 \leq y_1,$$
$$R_2(y_0, y_1) \equiv 2^{y_0} \leq y_1, \qquad R_3(y_0, y_1) \equiv 2_4^{y_0} \leq y_1$$

such that every set H which is homogeneous for all R_i: $u = 0, \ldots, 3$, and satisfies $|H| > [h_0/2]$ and $h_0 \geq 3$ is also 2_4^x-scattered, cf. [21, Lemma 2.13].

Now we encode $R(t, \bar{y})$ and $R_i(\bar{y})$ for $i = 0, \ldots, 3$ into one relation

$$R'(t, \bar{y}) \Leftrightarrow [t < \lg^2(y_0) \wedge R(t, \bar{y}) \vee \exists i \leq 3\, (t = \lg^2(y_0) + i \wedge R_i(\bar{y}))].$$

Observe that

(2) for every set $H \subseteq A$, $\lg^2(x) + 4$-diagonally homogeneous for $R'(t, \bar{y})$ and satisfying $|H| \geq [h_0/2]$ (with $h_0 \geq c \geq 3$), the set $\lg^4[H \setminus \{h_0\}]$ is an a-skeleton,

because then H is homogeneous for R_i: $i = 0, \ldots, 3$, i.e., it is 2_4^x-scattered and thus satisfies the assumptions of (1).

Finally, let $P = S_R \subseteq [A]^{2m+1}$ be the relation from Lemma 4.3 corresponding to R and to the function $F(x) = \lg^2(x) + 3$. We can assume that $P \subseteq [A]^a$. Let $H \in \mathrm{Hom}(P)$ and $|H| \geq h_0$. If c is large enough then for $h_0 \geq c$ we have $a(2^{(\lg^2 h_0)+4} + 1) < h_0$, because $a \leq \lg^4 h_0$. Hence by 4.3, the set H minus its last m elements is $\lg^4(x) + 3$-diagonally homogeneous for $R(t, \bar{y})$.

It follows that the set $H' = H$ minus its last a elements satisfies the assumption (2) and thus $\lg^4[H' \setminus \{h_0\}]$ is an a-skeleton. Obviously $\lg^4 H' = \lg^4 H$ minus its last a elements, which finishes the proof. □

Observe that also $\lg^4[H \setminus \{h_0\}]$ minus its last $\frac{1}{2}2_4^a$ elements is an a-skeleton.

4.5. Corollary. *There exists a term number c such that $\mathrm{I}\Delta_0 + \exp$ proves that to every set A and to every a we can associate $P : [A]^2 \to 2$ such that if $H \in \mathrm{Hom}(P)$, H is relatively large and $|H| \geq a + \max(2_4^a, c)$, then $h_0 \geq \max(2_4^a, c)$ and the set $\lg^4[H \setminus \{h_0\}]$ minus its last a elements is an a-skeleton.*

Proof. Observe first that Lemma 4.2 will be valid if we weaken the assumption $|H| \geq h_0$ to $|H| \geq h_0 - 1$ and if we change a to $a - 1$. Let c be the constant from the new version of Lemma 4.2 and let $Q : [A]^{a-1} \to 2$ be a relation as in that lemma. Let $e = \max(2_4^a, c)$.

We define $P(u_1, \ldots, u_a)$ to equal $Q(u_1, \ldots, u_{a-1})$ if $u_1 \geqslant e$, and $1 - Q(u_2, \ldots, u_a)$ if $u_1 < e$. Assume that $H \in \mathrm{Hom}(P)$, H is relatively large and $|H| \geqslant a + e$. We take $i < e$ such that $h_i < e \leqslant h_{i+1}$. Hence $i + a \leqslant e + a - 1$ and $P(h_i, \ldots, h_{i+a-1}) = 1 - Q(h_{i+1}, \ldots, h_{i+a-1})$ but $P(h_{i+1}, \ldots, h_{i+a}) = Q(h_{i+1}, \ldots, h_{i+a-1})$, which is a contradiction, because $H \in \mathrm{Hom}(P)$. Thus $h_0 \geqslant e = \max(2_4^a, c)$. Moreover, obviously $h \setminus \{\max H\}$ is homogeneous for Q and by Lemma 4.2 we obtain the desired conclusion.

4.6. Note. The above corollary shows that Lemma I.1.8 (on the existence of a-skeletons for all a) is also valid if the assumption PH is weakened to

$$\forall e\, \forall k \exists M\, \forall (P : [M]^e \to 2)\, \exists H \in \mathrm{Hom}(P)\, (|H| \geqslant \max(h_0, k)).$$

In the sequel by PH, we rather mean the above principle, which is equivalent in $\mathrm{I}\Delta_0 + \exp$ to the original formulation; using II.1.13 one can easily show that this principle is equivalent to $\mathrm{R}(\mathrm{PA}; \Sigma_1)$.

In [25] it was proved that the syntactically weakest variant of PH (with $k = e + 1$) is also equivalent to PH.

III. Arithmetical transfinite induction and Hardy hierarchies

In this chapter we collect all results which are connected with Wainer's theorem on provably recursive functions. We generalize this theorem to $\mathrm{I}(\varepsilon_\alpha)$. Originally it was proved for PA and says that if f is provably recursive in PA then there exists an $\alpha < \varepsilon_0$ such that $\exists y\, \forall x \geqslant y\, f(x) < H_\alpha(x)$.

III.1. Main results

All results referring to provably total functions in $\mathrm{I}(\varepsilon_\alpha)$, in this chapter, are in fact consequences of one more general theorem. To formulate it we need three notions: of provably total function in $\mathrm{I}(\varepsilon_\alpha)$, of $(\mathrm{B}^+)^2$-ε-systems of sequences and the notion of theories of Hardy hierarchies.

1.1. We say that a function $f : \mathbb{N} \to \mathbb{N}$ is provably total in a theory T and is of class Σ_n in T iff there exists $\varphi \in \Sigma_n$ such that φ defines f in $\mathbb{N}$ and $T \vdash \forall x\, \exists!\, y\, \varphi(x, y)$. We will write shortly: f is total and Σ_n in T.

Functions provably recursive in T are functions which are total and Σ_1 in T. The basic provably recursive function is the successor; however, the basic provably total function of class Σ_{n+1} (for $n \geqslant 1$) is the function H^n (see Definition II.2.9), the smallest majorant of all total functions of class Σ_n. This fact is provable in $\mathrm{I}\Sigma_n$ and H^n may be treated as basic in all theories extending $\mathrm{I}\Sigma_n$. For $n = 0$ we make the convention $H^0(x) = x + 1$.

Let $\bar{H}^n$ denote the function defined by H^n in the model $\mathbb{N}$. It follows that for every $k \in \omega$ the usual iteration $(\bar{H}^n)^k$ is provably total and Σ_{n+1} in $\mathrm{I}\Sigma_n$, for $n \geqslant 1$;

and moreover, that every function which is provably total and Σ_{n+1} in $\mathrm{I}\Sigma_n$ is majorized by some $(\bar{H}^n)^k$. Since $\mathrm{PA} \equiv \mathrm{I}(<\varepsilon_0)$, $(\bar{H}^n)_\alpha$ is provably total and Σ_{n+1} in PA for each $\alpha < \varepsilon_0$. Moreover, one can prove that each such function is majorized by a certain $(\bar{H}^n)_\alpha$, where $\alpha < \varepsilon_0$ (implicit in [20] and in [23]).

1.2. Assume now that

(*) λ together with accompanying formulas defines an ε-system of notations of class Σ_1 and let $P(\alpha, n)$ be an arbitrary Σ_1 formula.

We say that a system P of sequences for λ is a $(\mathrm{B}^+)^2$-ε-system of sequences if

(a) P is a B^+-ε-system,

(b) there exists a 'B^+-like mapping' $Q(\alpha, n) = \beta \in \Sigma_1$, i.e., $Q(\alpha, n) + 1 < \beta < \alpha$ implies $Q(\alpha, n) + 1 \leqslant Q(\beta, 0)$, such that

(c) $(\varepsilon_\alpha)_n^P = \varepsilon_{\alpha_n^Q}$ for $\alpha \in Lim \wedge \alpha > 0 \wedge n > 0$.

At the same time observe that by I.2.1 and I.2.2 it follows that for every system λ of notation in $\mathrm{I}\Sigma_1$ of class Σ_1 there exists a $P \in \Sigma_1$ which defines in $\mathrm{I}\Sigma_1$ a $(\mathrm{B}^+)^2$-ε-system of sequences for λ.

1.3. Let $n \in \mathbb{N}$ and let $\alpha < \lambda$. P-$\mathrm{TH}^n(<\varepsilon_{\alpha+1})$ denotes the theory in the language L_{PA} having axioms $\mathrm{I}\Sigma_1 \restriction \Pi_2$ (cf. Introduction) plus sentences stating that "P is a $(\mathrm{B}^+)^2$-ε-system of sequences for $\varepsilon_{\alpha+1}$" plus all the sentences of the form $\forall x\, \exists y\, (H^n)_\beta(x) = y$ for β of the form $\omega_m^{\varepsilon_\alpha+1}$, where m varies over the natural numbers.

Now we can formulate the most general result about functions provably total in $\mathrm{I}(\varepsilon_\alpha)$, in connection with the width of the proof. The assumption (*) will be a standing assumption in this section.

1.4. Theorem. *There exist constants c, c_1 such that for every $\varepsilon_\alpha < \lambda$ and for each φ of class Π_n and for each $m \in \mathbb{N}$,*

$$\mathrm{I}(\varepsilon_\alpha) \vdash_m \forall x\, \exists y\, \varphi(x, y)$$

implies

$$P\text{-}\mathrm{TH}^n(<\varepsilon_{\alpha+1}) \vdash_{m^{c_1}} \forall x \geqslant \underline{m}^c\, \exists y < (H^n)^P_{\beta(m,c)}(x)\, \varphi(x, y) \wedge$$
$$\forall x < \underline{m}^c\, \exists y < (H^n)^P_{\beta(m,c)}(\underline{m}^c)\, \varphi(x, y), \quad \text{where } \beta(m, c) = \omega_{m^c}^{\varepsilon_\alpha+1}.$$

If $\varepsilon_{\alpha+1}$ is not defined then we make the convention that $\varepsilon_{\alpha+1}$ denotes the standard expansion of ε_α.

The proof of the above theorem is based on a lemma, which we now formulate, but postpone its rather long proof to the next section. In the formulation we make the convention that

$$\varepsilon_{-1} = \omega, \qquad a^{1/c} = \max_b b^c \leqslant a.$$

1.5. Lemma. *There exists a term number c such that $I\Sigma_1 +$ "P defines a $(B^+)^2$-ε-system of sequences for λ" proves that for all α satisfying $\alpha = -1 \vee \alpha < \lambda$, and for each $\varepsilon_\alpha + 1$-P-large a-skeleton A,*

$$A \vDash_{a^{1/c}} \mathrm{Ax}(\mathrm{I}(\varepsilon_\alpha)).$$

1.6. Note. It is an open problem whether Lemma 1.5 and Theorem 1.4 are true for B^+-ε-systems. But observe that $(\mathrm{B}^+)^2$-ε-systems are still very natural. For example there exists a natural B^+-ε-system for the Feferman number Γ_0 having the property that $(\varepsilon_\alpha)_n = \varepsilon_{\alpha_{n-1}}$ for $n > 0$ and $\alpha \neq \varepsilon_\alpha$ (cf. Schmidt [28]). It is obvious that such a system is also a $(\mathrm{B}^+)^2$-ε-system with Q such that $\alpha_n^Q = \alpha_n^P$ when $(\alpha = \varepsilon_\alpha \vee n = 0)$ and $\alpha_n^Q = \alpha_{n-1}^P$ when $\alpha \neq \varepsilon_\alpha \wedge n > 0$.

Moreover, observe the following. Assume that $(\varepsilon_\alpha)_n^P \Rightarrow_0 0 \wedge n > 0 \wedge \alpha > 0$. Since $\varepsilon_{\beta+1} \Rightarrow_0 \varepsilon_\beta$ for each β, $\varepsilon_{\alpha_n^Q} \Rightarrow_0 0$ implies $\alpha_n^Q \Rightarrow_0^Q 0$. But Q is a B^+-system, hence $\alpha_n^Q \Rightarrow_0^Q \alpha_{n-1}^Q + 1$. Finally $(\varepsilon_\alpha)_n^P \Rightarrow_0 \varepsilon_{\alpha_{n-1}^Q + 1}$. Putting $\alpha_n^{Q'} = \alpha_{n-1}^Q$ we conclude that

(a) there exists a $Q' \in \Sigma_1$ such that $\sup \alpha_n^{Q'} = \alpha$ for each $\alpha \in Lim \setminus \{0\}$ and $(\varepsilon_\alpha)_n^P \Rightarrow_0 0$ implies $(\varepsilon_\alpha)_n^P \Rightarrow_0 \varepsilon_{\alpha_n^{Q'}+1}$ for $\alpha \in Lim \setminus \{0\}$.

All the above reasoning is formalizable in $\mathrm{I}\Sigma_1 +$ "P is a $(\mathrm{B}^+)^2$-ε-system".

In the proof of 1.5 we in fact use the property (a) of $(\mathrm{B}^+)^2$-ε-systems. We will see that all results are valid for B^+-ε-systems having the property (a). But we explicitly assume the more elegant property that P is a $(\mathrm{B}^+)^2$-ε-system.

Proof of Theorem 1.4. We take for c the constant from Lemma 1.5. We fix an α such that $\varepsilon_\alpha < \lambda$ and assume that $\mathrm{I}(\varepsilon_\alpha) \vdash_m \forall x\, \exists y\, \varphi(x, y)$, where φ is of class Π_n. By Σ_1-completeness

$$\mathrm{I}\Sigma_1 \vdash (\mathrm{I}(\varepsilon_\alpha) \vdash_{\underline{m}} \ulcorner \forall x\, \exists y\, \varphi(x, y) \urcorner).$$

Now we work in the system P-$\mathrm{TH}^n(<\varepsilon_{\alpha+1})$. First we construct an informal proof. The length of a formal proof will be estimated at the end.

Take an $a \geqslant \underline{m}^c$. Let $d = (H^n)^P_{\beta(m,c)}(a)$. Let $f = H^n \cap [a, d]^2$. We define $A = \{f^l(a) : f^l(a) \leqslant d\}$. Hence by II.3.2, the set A is $\beta(m, c)$, P-large and H^n-scattered (i.e., Σ_n-scattered).

By II.3.4 there exists an $\varepsilon_\alpha + 1$, P-large $\underline{m}^c$-skeleton $B \subseteq A$. In fact we need an obvious strengthening of II.3.2 and II.3.4 with $\mathrm{I}\Sigma_1$ replaced by $\mathrm{I}\Sigma_1 +$ "P defines a B^+-ε-system of sequences for λ".

By Lemma 1.5, $B \vDash_{\underline{m}} \mathrm{Ax}(\mathrm{I}(\varepsilon_\alpha))$. It follows by the Soundness Lemma II.1.10 that $B \vDash \ulcorner \forall x\, \exists y\, \varphi(x, y \urcorner$. Hence (cf. II.1.9), $\forall x < b_0\, \exists y < b_1\, B \vDash \ulcorner \varphi \urcorner (x, y)$; we can assume that c is large enough to guarantee us that $\ulcorner \varphi \urcorner \neg \leqslant \underline{m}^c$. Thus, by the Absoluteness Lemma II.2.4, $\forall x < b_0\, \exists y < b_1\, \mathrm{Tr}_{\Pi_n}(\ulcorner \varphi \urcorner (x, y))$. Therefore $\forall x < b_0\, \exists y < b_1\, \varphi(x, y)$, which implies that $\forall x \leqslant a\, \exists y < (H^n)^P_{\beta(m,c)}(x)\, \varphi(x, y)$, because $b_1 \leqslant d$. Taking $x = a$ we obtain the first conjunct of the statement of the lemma. Taking at the beginning $a = \underline{m}^c$ we obtain the second component. This finishes an informal proof.

Before we start an estimation of the width of a formal proof, we observe that without loss of generality we can assume that in the given proof of the formula $\forall x\, \exists y\, \varphi(x, y)$ in $\mathrm{I}(\varepsilon_\alpha)$ an axiom of the form $\mathrm{Ind}(\psi, \varepsilon_{\underline{\alpha}})$ was used at least once. This guarantees that the code of $\underline{\alpha}$ is less than m. In the opposite case we have $\mathrm{I}(\varepsilon_{-1}) \vdash_m \forall x\, \exists y\, \varphi(x, y)$ and we can assume that $\alpha = -1$.

Now for the direct estimation of width observe that all steps in the proof, where we refer to the proven lemmas (1.5, II.3.2, II.3.4, II.1.10, II.1.9, II.2.4) have in total a fixed (independent of m) length c_1.

All the steps in the proof in P-$\mathrm{TH}^n(<\varepsilon_{\alpha+1})$, which we make directly after the steps listed above consist in substitution of some terms, e.g. $\underline{\alpha}$ to Lemma 1.5, $\underline{m}^c$ and $\underline{\alpha}$ to Lemma II.3.4, $\ulcorner \forall x\, \exists y\, \varphi \urcorner$ and $\underline{m}$ to the Soundness Lemma, $\ulcorner \varphi \urcorner$ to the Absoluteness Lemma, etc. The codes of these terms are $\leqslant m^{c_2}$, where c_2 is a constant independent of m. Hence the codes of all formulas which are results of some substitutions are $\leqslant m^{c_3}$, where c_3 is independent of m (cf. II.2.12 and the proof of II.2.13).

Let us now estimate the width of the step which we make at the very beginning. Let $p = \varphi_1, \varphi_2, \ldots, \varphi_l$ be a proof of $\forall x\, \exists y\, \varphi(x, y)$ in $\mathrm{I}(\varepsilon_\alpha)$ having width $\leqslant m$. Then the sequence $\ulcorner \varphi_1 \urcorner, \ulcorner \varphi_2 \urcorner, \ldots, \ulcorner \varphi_l \urcorner$ is a formal proof (of width $\leqslant \underline{m}$) from $\mathrm{I}(\varepsilon_{\underline{\alpha}})$ in the theory $\mathrm{I}\Sigma_1$. Since the formula $\mathrm{I}(\varepsilon_{\underline{\alpha}}) \vdash_{\underline{m}} \ulcorner \varphi_i \urcorner$ is built up of $\underline{\alpha}$, $\underline{m}$ and $\ulcorner \varphi_i \urcorner$ in a fixed number of steps (independent of m), it has a code $\leqslant m^{c_4}$, where c_4 is a constant (independent of m). Using this fact we can find a constant c_5 such that $\forall i \leqslant l\; \mathrm{I}\Sigma_1 \vdash_{m^{c_5}} (I(\varepsilon_{\underline{\alpha}}) \vdash_{\underline{m}} \ulcorner \varphi_i \urcorner)$, which is proved by induction on i.

The next step in the proof is a reference to the axiom $\forall x\, \exists y\, (H^n)^P_{\beta(m,c)}(x) = y$. This formula is built up, in a fixed number of steps, of $\underline{m}^c$, $\underline{\alpha}$ and H^n. The length of H^n is $\leqslant c_6 \cdot n$.

Since $2^n \leqslant \ulcorner \varphi \urcorner \leqslant m$, it follows that the code of H^n is $\leqslant m^{c_7}$. Hence finally the code of the axiom is $\leqslant m^{c_8}$.

All the remaining steps are logical inferences having a scheme independent of m. Hence the width can only grow polynomially. Thus we can construct a formal proof having width $\leqslant m^{c_0}$, where c_0 is a constant independent of m. $\square$

Now we deal with some corollaries. Assume that

(**) $P \in \Sigma_1$ defines in $\mathrm{I}\Sigma_1$ a $(\mathrm{B}^+)^2$-ε-system of sequences for $\varepsilon_{\alpha+1}$.

The following direct generalization of Wainer's Theorem is an immediate consequence of Theorem 1.4.

1.7. Corollary. *Assume that $\varepsilon_{\alpha+1}$ is a well-ordering. If f is provably total and Σ_n in $\mathrm{I}(\varepsilon_\alpha)$ then there exists a $\beta < \varepsilon_{\alpha+1}$ such that $\exists y\, \forall x \geqslant y\, f(x) < (H^n)^P_\beta(x)$.*

An easy refinement of Theorem 1.4 shows that the assumption (**) can be weakened to: P is a B-ε-system satisfying 1.6(a), not necessarily definable in the standard model $\mathbb{N}$. Since all H^P_β: $\beta < \varepsilon_{\alpha+1}$ are provably total in $\mathrm{I}(\varepsilon_\alpha)$, we have a simple corollary.

1.8. Corollary. $\mathrm{Rec}(\mathrm{I}(\varepsilon_\alpha)) = \bigcup_{\beta<\varepsilon_{\alpha+1}} E(H^P_\beta)$, *where* $E(f)$ *denotes the operation which associates with* f *the class of all elementary recursive functions from* f, *introduced by Grzegorczyk* [8] *to classify the primitive recursive functions.*

Proof. Assume that $f \in \mathrm{Rec}(\mathrm{I}(\varepsilon_\alpha))$. Hence there exists $\varphi(x, y, z) \in \Delta_0$ such that $\exists z\, \varphi(x, y, z)$ defines f in $\mathbb{N}$ and $\mathrm{I}(\varepsilon_\alpha) \vdash \forall x\, \exists! y\, \exists z\, \varphi(x, y, z)$. By 1.7 there exists $\beta < \varepsilon_{\alpha+1}$ such that $N \vDash \forall x\, \exists y, z \leq H_\beta(x)\, \varphi(x, y, z)$, which proves that $f \in E(H_\beta)$, because $E(H_\beta)$ is closed under bounded projections and bounded minimum. □

The next immediate consequence of 1.4, under the assumption (**), is:

1.9. Corollary. $\mathrm{I}(\varepsilon_\alpha) \restriction \Pi_{n+2} \equiv P\text{-}\mathrm{TH}^n(<\varepsilon_{\alpha+1})$.

In fact only the inclusion $\subseteq$ is immediate. For the proof of the opposite inclusion it is enough to recall that all the axioms of $P\text{-}\mathrm{TH}^n(<\varepsilon_{\alpha+1})$ are of class Π_{n+2}. Corollary 1.9 has a subtle variant, more close to 1.4, which concerns a 'polynomial reduction' of one theory to another.

1.10. Corollary.

$$\exists c\, \forall m\, \forall \varphi \in \Pi_{n+2}\, (\mathrm{I}(\varepsilon_\alpha) \vdash_m \varphi \rightarrow P\text{-}\mathrm{TH}^n(<\varepsilon_{\alpha+1}) \vdash_{m^c} \varphi).$$

Symbolically: $\mathrm{I}(\varepsilon_\alpha) \subseteq^{\mathrm{pol}}_{\Pi_{n+2}} P\text{-}\mathrm{TH}^n(<\varepsilon_{\alpha+1})$.

This is an obvious consequence of 1.4. The same is true for lengths of proofs. Moreover, it is obvious that there exists a primitive recursive map which sends proofs of width m (for $\varphi \in \Pi_{n+2}$) in one theory to proofs of width $\leq m^c$ in the second theory. It is not known whether the converse reduction is polynomial. A rough estimate of the upper bound is $2^{2^{c \cdot m}}$ (cf. (4) in the proof of the main Lemma in Section III.2).

All combinatorial-logical proofs presented in the present paper provide this sort of primitive recursive reduction, but we do not develop this point further. In this respect the combinatorial-logical proofs are similar to proof-theoretic proofs of reductions (see discussion on this subject in Feferman [6]), but additionally we have information about width and length of proofs.

1.11. Note. Now assume that $P \in \Sigma_1$ defines in $\mathrm{I}(\varepsilon_\alpha)$ a $(\mathrm{B}^+)^2$-ε-system of sequences for $\varepsilon_{\alpha+1}$ and that $\varepsilon_{\alpha+1}$ is well-ordered. Then it follows by 1.9($\supseteq$) and strengthened 1.7 that the hierarchy H^P_β: $\beta < \varepsilon_{\alpha+1}$ is majorized by every H^Q_β: $\beta < \varepsilon_{\alpha+1}$, where Q is a $(\mathrm{B}^+)^2$-ε-system, i.e. P is in a sense minimal.

It is not true that every $(\mathrm{B}^+)^2$-ε-system P is minimal. For each $f: \mathbb{N} \rightarrow \mathbb{N}$ one can construct a system P on $\varepsilon_\omega + 1$ such that $H^P_{\varepsilon_\omega}(n) \geq f(n)$ for all n. Hence if f majorizes a hierarchy of length $\varepsilon_{\alpha+1}$ (for $\alpha \geq \omega$), then no extension of P to $\varepsilon_{\alpha+1}$ is minimal. The construction of P is very technical but not difficult and we omit it.

Still some easy corollaries of Lemma 1.5 are worth stating. Assume that $\varepsilon_\alpha < \lambda$.

1.12. Corollary. *Let $n \geqslant 1$. $\mathrm{I}\Sigma_1$ proves the following equivalencies:*

$$\mathrm{R}(\mathrm{I}(\varepsilon_\alpha);\Sigma_n) \Leftrightarrow \forall a\, \exists A \text{ “}A \text{ is } \varepsilon_\alpha\text{-large } a\text{-skeleton} \wedge A \text{ is } \Sigma_{n-1}\text{-scattered”} \Leftrightarrow \mathrm{CR}(\mathrm{I}(\varepsilon_\alpha);\Sigma_n).$$

Proof (in $\mathrm{I}\Sigma_1$). Let us call (1), (2), (3) the above three sentences. Corollary II.3.6 says that (1)$\Rightarrow$(2). By Lemma 1.5, (3) immediately follows from (2), and (3)$\Rightarrow$(1) is the content of Lemma II.2.6. □

The next corollary is the Kreisel–Levy Theorem [14].

1.13. Corollary. $\mathrm{I}(\varepsilon_{\alpha+1}) \vdash \mathrm{R}(\mathrm{I}(\varepsilon_\alpha);\Sigma_n)$.

Proof. Corollary II.3.7 reads $\mathrm{I}(\varepsilon_{\alpha+1}) \vdash (2)$, where (2) comes from the proof of Corollary 1.12. By 1.12, the assertion follows. □

III.2. Proof of the Main Lemma III.1.5

First we prove the following generalization of Lemma II.3.4 for $(\mathrm{B}^+)^2$-ε-systems. (We make the convention that if $\alpha = 0$, then $\alpha_x = -1$.)

2.1. Lemma. *$I\Sigma_1$ + “P is a $(B^+)^2$-ε-system” proves that for each $\alpha < \lambda$, if a set A is ε_α, P-large then for each $z < a_0$ there exists a $B \subseteq A$ which is an $\varepsilon_{\alpha_{\varrho'}} + 1$, P-large a_0-skeleton, where Q' is such that $(\varepsilon_\alpha)_z^P \Rightarrow_0 \varepsilon_{\alpha_{\varrho'}+1}$ for $\alpha \in Lim$ (cf. Note* III.1.6(a)).

Proof. We have three cases.

Case (i): $\alpha = 0$. Then by II.3.4(i) there exists an a_0-skeleton $B \subseteq A$.

Case (ii): $\alpha = \beta + 1$. Then by II.3.4(ii) there exists an $\varepsilon_\beta + 1$, P-large a_0-skeleton B included in A. Obviously B is $\varepsilon_{\alpha_{\varrho'}} + 1$, P-large for $z < a_0$.

Case (iii): $\alpha \in Lim$. Then A is $(\varepsilon_\alpha)_{a_0}$, P-large. By I.2.6, $(\varepsilon_\alpha)_{a_0} \Rightarrow_0 0$ and we also have $(\varepsilon_\alpha)_{a_0} \Rightarrow_0 (\varepsilon_\alpha)_z = \varepsilon_{\alpha_{\varrho'}+1}$ for $z \leqslant a_0$. Therefore for each $z \leqslant a_0$, A is $\varepsilon_{\alpha_{\varrho'}+1}$, P-large and arguing as in (ii) we find an $\varepsilon_{\alpha_{\varrho'}} + 1$, P-large a_0-skeleton $B \subseteq A$. □

In the proof of Lemma III.1.5 we use all conventions described in II.2.12, especially those which refer to the use of the phrase ‘concrete numbers’.

We will work in $\mathrm{I}\Sigma_1$ + “P defines a $(\mathrm{B}^+)^2$-ε-system”.

We fix an interval $[0, b]$ and prove that there exists a concrete constant c_0 (independent of b) such that the conclusion of the lemma holds under the additional assumption that $A \subseteq [0, b]$. We will put the constant c_0 equal to the maximum of some chosen constants c_1, c_2, c_3, which guarantee the validity of successive steps in the proof. By Lemma I.2.6 on finitization the set of all α satisfying the condition $\exists A \subseteq [0, b]$ “A is a ε_α, P-large” is finite and we can argue freely by induction on α.

The initial step for $\alpha = -1$ follows for $c \geqslant c_1$, where c_1 is the constant such that $c_1 \geqslant \theta$ for $\theta \in \mathrm{Ax}(\mathrm{PA}^-)$. Indeed, if $a^{1/c} < 2$ then there is nothing to prove; but otherwise $a \geqslant 2^c \geqslant c_1$, hence $A \vDash \mathrm{Ax}(\mathrm{PA}^-)$ and by Lemma II.1.12 $A \vDash \mathrm{Ax}(\mathrm{PA})$.

Now we show the inductive step from all $\beta < \alpha$ to α, for $\alpha \geq 0$. Let $A \subseteq [0, b]$ be an $\varepsilon_\alpha + 1$, P-large a-skeleton. Additionally, assume that $\mathrm{Ind}(\theta, \varepsilon_\alpha) < a^{1/c}$. Our aim is to show that $A \models \mathrm{Ind}(\theta, \varepsilon_\alpha)$ provided c is greater than a concrete constant $c_0 \geq c_1$. Then the lemma will be proved. It follows from the assumption that θ^c, $\ulcorner\alpha\urcorner^c < a$. In the sequel instead of ω_x^β we write $\omega_x(\beta)$.

The formula $\mathrm{Ind}(\varepsilon_\alpha)$ is obviously equivalent in $\mathrm{I}\Sigma_1$ to

$$\forall k\, \forall z\, \forall \beta\, [\beta = \underset{\sim}{\alpha}_z^{Q'} \rightarrow \mathrm{Ind}(\theta, \omega_k(\varepsilon_\beta + 1)].$$

The usual Hilbert type proof (of the above-mentioned equivalence) has a concrete scheme independent of θ and $\underset{\sim}{\alpha}$, i.e., it is of the form $\psi_1(\theta), \ldots, \psi_{r-2}(\theta)$, $\forall y\, \psi_{r-1}(\theta, y)$, $\psi_r(\theta, \underset{\sim}{\alpha})$, where $\psi_1, \ldots, \psi_r$ are concrete formulas of $L_{\mathrm{PA}}(R)$, and r is a concrete number.

It is clear that there exists a concrete number c_2 (see II.2.12(3)) such that the width of this proof is less than $\max(\theta, \ulcorner\alpha\urcorner)^{c_2}$ (we can assume that $\theta \geq 2$).

To obtain a Hilbert style proof with axioms of substitutions of only simple terms as in Section II.1 we should proceed the above sequence by a sequence of axioms of substitutions of successive terms and modus ponens rules (cf. the proof of II.2.13). Hence also for this changed proof we can assume that its width is less than $\max(\theta, \ulcorner\alpha\urcorner)^{c_2}$, for suitably chosen concrete c_2 (we can assume that $c_2 \geq c_1$). Hence if $c \geq c_2$ from the very beginning then the considered width is less than a.

Thus by the Soundness Lemma II.1.10 to prove $A \vdash \mathrm{Ind}(\theta, \varepsilon_\alpha)$ it is enough to prove that

$$A \models \forall k\, \forall z\, \forall \beta\, [\beta = \underset{\sim}{\alpha}_z^{Q'} \rightarrow \mathrm{Ind}(\theta, \omega_k(\varepsilon_\beta + 1)].$$

By II.19 this reduces to showing that

$$\forall k < a_0\, \forall z < a_0\, \forall \ulcorner\beta\urcorner < a_0 [A \models \beta = \underset{\sim}{\alpha}_z^{Q'} \rightarrow A \models \mathrm{Ind}(\theta, \omega_k(\varepsilon_\beta + 1)].$$

We fix $k < a_0$, $z < a_0$ and $\ulcorner\beta\urcorner < a_0$ such that $A \models \beta = \underset{\sim}{\alpha}_z^{Q'}$. Since Σ_1 formulas are upwards absolute (see II.2.14), $\beta = \underset{\sim}{\alpha}_z^{Q'}$. By Lemma 2.1 there exists an $\varepsilon_{\alpha_z^{Q'}} + 1$, P-large a_1-skeleton $B \subseteq A \setminus \{a_0\}$.

We show that $B \models \mathrm{Ind}(\theta, \omega_k(\varepsilon_\beta + 1))$. This is enough to finish the proof, because A is diagonally indiscernible for the 'stars' of formulas less than a and we obtain $A \models \mathrm{Ind}(\theta, \omega_k(\varepsilon_\beta + 1))$.

By the inductive assumption

(1) $\quad B \models_{(a_1)^{1/c}} \mathrm{Ax}(\mathrm{I}(\varepsilon_\beta))$.

Let $m < a$ be such that $\theta \in \Pi_m$. Now we use the following fact: Π_m-induction up to $\omega_k^{\varepsilon_\beta + 1}$ reduces to the Π_{m+k}-induction up to ε_β—this is an old result of Gentzen (for the proof see [10]).

More precisely, the required fact is the following:

$$\forall \theta \in \Pi_m\, \exists \pi \in \Pi_{m+k}\, \mathrm{PA} + \mathrm{Ind}(\pi, \varepsilon_\beta) \vdash \mathrm{Ind}(\theta, \omega_k^{\varepsilon_\beta + 1}).$$

The proof of this fact consists in a simple application of the following reduction step: for all terms $\underline{\gamma}$ such that $\mathrm{PA} \vdash \underline{\gamma} < \lambda$,

(2) $\quad \forall n\, \forall \theta' \in \Pi_n\, \exists \eta' \in \Pi_{n+1}\ \mathrm{PA} + \mathrm{Ind}(\eta', \gamma) \vdash \mathrm{Ind}(\theta', \omega^{\underline{\gamma}})$.

We do not need to remember the proof of the reduction step. The only important information for us is that the reduction formula η' is a concrete expression built-up from θ' and γ, i.e., it has the form $\psi(\theta', \underline{\gamma})$. If $\underline{\gamma} = \omega_{\underline{l}}^{\varepsilon_\beta + 1}$ then η' is a concrete expression built-up from θ', $\underline{\beta}$, $\underline{l}$ and symbols of L_{PA}.

Also the proof of the reduction has the concrete character. Namely, it is a concrete sequence $\psi_1(\theta', \underline{\beta}, \underline{l}), \ldots, \psi_r(\theta', \underline{\beta}, \underline{l})$. We can assume that this sequence includes the proof of the equality $\omega^{\underline{\gamma}} = \omega_{l+1}(\varepsilon_\beta + 1)$. Hence there exists a concrete constant d such that for all n, l ($l > 0$)

(3) $\quad \forall \theta' \in \Pi_n\, \exists \eta' \in \Pi_{n+1}\ \mathrm{PA} \cup \{\mathrm{Ind}(\eta', \omega_{\underline{l}}(\varepsilon_{\underline{\beta}} + 1))\} \vdash_{f(\theta', l)} \mathrm{Ind}(\theta' \omega_{l+1}(\varepsilon_{\underline{\beta}} + 1))$,

where $f(\theta', l) \leqslant \max(\theta', \ulcorner\beta\urcorner, l)^d$.

(4) **Claim.** *There exists a concrete constant c_3 such that for every a-skeleton A satisfying $a \geqslant c \geqslant c_3$, there exists an η such that*

$$\mathrm{Ax(PA)} \cup \{\mathrm{Ind}(\eta, \varepsilon_{\underline{\beta}})\} \vdash_b \mathrm{Ind}(\theta, \omega_{\underline{k}}(\varepsilon_{\underline{\beta}} + 1)), \quad \textit{where } b = (a_1)^{1/c}.$$

We define inductively

$$f'(\theta, 0) = f(\theta, 0), \qquad f'(\theta, l+1) = f(f'(\theta, l), l+1).$$

If we apply k times the reduction step (3) we obtain a formula $\eta \in \Pi_{m+k}$ such that

(4) $\quad \mathrm{Ax(PA)} \cup \{\mathrm{Ind}(\eta, \varepsilon_{\underline{\beta}})\} \vdash_{f'(\theta, k)} \mathrm{Ind}(\theta, \omega_{\underline{k}}(\varepsilon_\beta + 1))$.

Since $\theta < a^{1/c}$ and $\ulcorner\beta\urcorner < a_0$ it follows that $f'(\theta, 0) < a_0^d$ and inductively that $f'(\theta, k) < a_0^{d^{k+1}} \leqslant a_0^{d^{a_0}}$.

By II.2.11 there exists a concrete number c_3 such that every a-skeleton A with $a \geqslant c_3$ is $(x^{d^x})^x$-scattered apart a few final elements. Hence if $a \geqslant c \geqslant c_3$ then $(a_0^{d^{a_0}})^c < a_1$ and the claim is proved.

From Claim (4) it follows by (1) and the Soundness Lemma that $B \models \mathrm{Ind}(\theta, \omega_{\underline{k}}(\varepsilon_\beta + 1))$ if we assume from the very beginning that $c \geqslant c_0 = \max(c_2, c_3)$, i.e., the inductive step follows and the lemma is proved. □

2.2. Note. At the end I should like to point a direction of the possible refinement of the results of this chapter.

Let $\mathrm{I}(\Pi_n; \varepsilon_\alpha)$ denote the theory of $\mathrm{I}(\varepsilon_\alpha)$ with ε_α-induction restricted to the Π_n formulas. One can easily show that

$$\mathrm{I}(\Pi_n; \omega^{\varepsilon_\alpha \cdot k}) \subseteq \mathrm{I}(\Pi_n; \varepsilon_\alpha) \quad \text{for } k \in \omega.$$

Using (2) from the proof of 1.5 we obtain

$$\mathrm{I}(\Pi_2; <\omega_{n+1}^{\varepsilon_\alpha + 1}) \subseteq \mathrm{I}(\Pi_n; \varepsilon_\alpha) \quad \text{for } n \geqslant 2.$$

Hence

$$P\text{-TH}^0(<\omega_{n+1}^{\varepsilon_\alpha+1}) \subseteq \mathrm{I}(\Pi_n; \varepsilon_\alpha) \quad \text{for } n \geq 2.$$

One can also show the opposite inclusion. We only sketch the proof. Assume that $\mathrm{I}(\Pi_n; \varepsilon_\alpha) \vdash \forall x\, \exists y\, \varphi(x, y)$, where $\varphi \in \Delta_0$. Using the cut-free-cut elimination theorem (see [31] and [3, 4.3]), which is also valid for $\mathrm{I}(\Pi_n; \varepsilon_\alpha)$, we find a proof p of $\forall x\, \exists y\, \varphi(x, y)$ consisting only of Π_{n+2} formulas; let us stress that we can descend to Π_{n+1} formulas, if we use rules of induction.

Working in $P\text{-TH}^0(<\omega_{n+1}^{\varepsilon_\alpha+1})$ let $\beta = \omega_n^{\varepsilon_\alpha(m^c+2)}$ and let $X = [m^c, H_\beta(m^c)]$; c will be chosen in the sequel. Then X is $X = [m^c, H_\beta(m^c)]$; c will be chosen in the sequel. Then X is $\omega_n^{\varepsilon_\alpha(x_0+2)}$-large and by II.3.13(b) we can find an ε_α-large set A such that $\lg^3[A]$ is an $(n+1, m^c)$-skeleton.

If θ is a sentence of the form $\forall \bar{x}\, \eta(\bar{x})$, where η is a boolean combination of Σ_{n+1} formulas ($\eta \in B(\Sigma_{n+1})$) then we define

$$A \vDash \theta \Leftrightarrow \forall \bar{x} < a_0\, A \setminus \{A_0\} \vDash \eta(\bar{x}).$$

One can check that the Soundness Lemma is valid for $A \vDash \theta$ if we restrict proofs to proofs consisting of Π_{n+2} formulas.

Hence what remains is to show that $A \vDash \mathrm{Ind}(\theta, \varepsilon_\alpha)$ for $\theta < m \wedge \text{a.r.}(\theta) \leq n$. We write $\mathrm{Ind}(\theta, \varepsilon_\alpha)$ in the form

$$\forall k\, \forall z\, \forall \beta\, \forall a\, [\beta = \alpha_z^{Q'} \to \mathrm{Ind}'(\theta(a), \omega_k(\varepsilon_\beta + 1))],$$

where Ind′ indicates that we do not use parameters. Then as in the proof of 1.5 we fix k, z, $\ulcorner\beta\urcorner$, $b < a_0$ and show that $B \vDash \mathrm{Ind}'(\theta(b), \omega_k(\varepsilon_\beta + 1))$ for a certain $B \subseteq A \setminus \{a_0\}$, if c is large enough. Since $\mathrm{Ind}'(\theta(b), \omega_k(\varepsilon_\beta + 1))$ is a $B(\Sigma_{n+1})$ formula it follows also that $A \setminus \{a_0\} \vDash \mathrm{Ind}'(\theta(b), \omega_k(\varepsilon_\beta + 1))$, which finishes the proof. □

IV. The α-time iterated principles

In this chapter we investigate the relation between the theory of the iterated reflection principle $\mathcal{R}_\alpha(\mathrm{PA})$ and the theory $\mathrm{I}(\varepsilon_\alpha)$ for all recursive ordinals α. Moreover, we consider the sentences PH^α (in the style of Paris and Harrington) and show that they are independent of $\mathrm{I}(<\varepsilon_\alpha)$.

IV.1. The iterations of the reflection principle

Assume that λ defines a basic system of notations in $\mathrm{I}\Sigma_1$. Then $\mathcal{R}_\alpha(\mathrm{PA})$: $\alpha < \lambda$ denotes any Σ_1-definable sequence of theories for which $\mathrm{I}\Sigma_1$ proves:

$$\mathcal{R}_0(\mathrm{PA}) \equiv \mathrm{PA} + \mathcal{R}(\mathrm{PA}),$$

$$\mathcal{R}_\alpha(\mathrm{PA}) \equiv \mathrm{PA} + \mathcal{R}\Big(\bigcup_{\beta<\alpha} \mathcal{R}_\beta(\mathrm{PA})\Big) \quad \text{for } \alpha \neq 0;$$

in fact, we add the formalized version of $\mathcal{R}(\bigcup_{\beta<\alpha}\mathcal{R}_\beta(\mathrm{PA}))$. The theory $\bigcup_{\beta<\alpha}\mathcal{R}_\beta(\mathrm{PA})$ will be denoted by $\mathcal{R}_{<\alpha}(\mathrm{PA})$. According to the definition of $\mathcal{R}(T)$ in II.2.5 the set $\mathcal{R}(\mathcal{R}_{<\alpha}(\mathrm{PA}))$ consists of the sentences:

$$\forall x\, (\ulcorner \exists y\, (\mathrm{Prf}(\mathcal{R}_{<\alpha}(\mathrm{PA}); y) \wedge \mathrm{end}(y) = \mathrm{Sub}(v, \underset{\sim}{x}))\urcorner (\underline{\eta}/v) \rightarrow \eta(x)),$$

where η varies over formulas of L_{PA}.

Here $\mathrm{Prf}(\mathcal{R}_{<\alpha}(\mathrm{PA}); y)$ means that y is a proof from $\mathcal{R}_{<\alpha}(\mathrm{PA})$, $\mathrm{end}(y)$ denotes the last term of the sequence y, $\mathrm{Sub}(v, \underset{\sim}{x})$ denotes $v(\underset{\sim}{x}/u)$, if $v \in L_{\mathrm{PA}}$ and u is the only free variable in v, $\ulcorner\varphi\urcorner$ denotes the term number mapped to the Gödel number of φ. For more details see [4] and [30].

Obviously $\mathrm{I}\Sigma_1$ proves that $\mathcal{R}_{<\alpha}(\mathrm{PA}) \subseteq \mathcal{R}_\alpha(\mathrm{PA})$ for $\alpha > 0$. Finally, we put $\mathrm{Ax}(\mathcal{R}_\alpha(\mathrm{PA})) := \mathrm{Ax}(\mathrm{PA}) \cup \mathcal{R}(\mathcal{R}_{<\alpha}(\mathrm{PA}))$.

In this section we present a new, combinatorial-logical, proof of the Kreisel–Levy–Schmerl theorem [14, 27].

1.1. Theorem. *If λ defines in $\mathrm{I}\Sigma_1$ and ε-system of notations, then for every $\alpha < \lambda$, $\mathrm{I}(\varepsilon_\alpha) \equiv \mathcal{R}_\alpha(\mathrm{PA})$.*

We only show the inclusion $\mathcal{R}_\alpha(\mathrm{PA}) \subseteq \mathrm{I}(\varepsilon_\alpha)$. The opposite inclusion was proved by Schmerl [27] by means of some tricky Gödel diagonal considerations; he used the Löb Theorem. This rarther simple, but useful trick also plays a role in the old proof-theoretical proof of the inclusion $\mathcal{R}_\alpha(\mathrm{PA}) \subseteq \mathrm{I}(\varepsilon_\alpha)$. Kreisel–Levy [14] proved the following reduction step: $\mathrm{I}(\varepsilon_\alpha) \vdash \mathcal{R}(\mathrm{I}(<\varepsilon_\alpha))$; but the rest of the proof of the inclusion $\mathcal{R}_\alpha(\mathrm{PA}) \subseteq \mathrm{I}(\varepsilon_\alpha)$ relies on the above-mentioned trick.

We present a combinatorial-logical proof of the 'complete reduction in one step', i.e., the inclusion $\mathcal{R}_\alpha(\mathrm{PA}) \subseteq \mathrm{I}(\varepsilon_\alpha)$. We do not use in this proof any form of diagonal considerations. In the companion paper [24], a combinatorial-logical proof of the reduction $\bigcup_n P\text{-}\mathrm{TH}^n(\varepsilon_\alpha) \vdash \mathcal{R}(\bigcup_n P\text{-}\mathrm{TH}^n(<\varepsilon_\alpha))$ is presented. This reduction together with the results of the previous chapter gives another proof of the Kreisel–Levy reduction step.

To formulate the main lemma for the proof of Theorem 1.1 assume that $P \in \Sigma_1$ defines in $\mathrm{I}\Sigma_1$ a $(\mathrm{B}^+)^2$-ε-system of sequences for λ; it exists by I.2.1 and I.2.2. For the definition see III.1.2.

We put $\varepsilon_{-1} = \omega$, $\mathrm{Ax}(\mathcal{R}_{-1}(\mathrm{PA})) = \mathrm{Ax}(\mathrm{PA})$.

1.2. Lemma. $\mathrm{I}\Sigma_1$ *proves: there exists a concrete number c such that for every $c \geqslant c_0$, and for every α satisfying $-1 \leqslant \alpha < \lambda$ and every a the following implication is valid:*

$$\text{“}A \textit{ is an } \varepsilon_\alpha + 1,\ P\textit{-large } a\textit{-skeleton”} \wedge b = (a)^{1/c} \rightarrow A \models_b \mathrm{Ax}(\mathcal{R}_\alpha(\mathrm{PA})).$$

Proof. We work in $\mathrm{I}\Sigma_1$. The underling idea of the proof is the same as the idea behind the proof of Lemma II.1.5.

We fix an interval $[0, d]$ and prove the lemma for $A \subseteq [0, d]$ by induction on α, which by Lemma I.2.6 on finitization reduces to finite induction. We chose successively some concrete constants c_1, c_2, c_3 for successive steps in the proof and put finally $c = \max(c_1, c_2, c_3)$ (cf. the beginning of the proof of III.1.5).

The initial step for $\alpha = -1$ follows from Lemma II.1.12 for $c \geq c_1$, where c_1 is a constant such that $c_1 \geq \theta$ for $\theta \in \mathrm{Ax}(\mathrm{PA}^-)$.

Now we show the inductive step for α in $\mathrm{I}\Sigma_1$. Let $A \subseteq [0, d]$ be an $\varepsilon_\alpha + 1$, P-large a-skeleton. Additionally we assume $\theta < b = a^{1/c}$, $\theta \in \mathrm{Ax}(\mathscr{R}_\alpha(\mathrm{PA}))$. Our aim is to show that $A \models \theta$ *provided* c is greater than a concrete constant $c_2 \geq c_1$ (independent of d), in all assumptions, also in the inductive assumption.

If $\theta \in \mathrm{Ax}(\mathrm{PA})$ then, by I.1.12, $A \models \theta$. Therefore to end the proof we have to consider the following case:

$$\theta := \forall x\, (\ulcorner \exists y\, (\mathrm{Prf}(\mathscr{R}_{<\alpha}(\mathrm{PA}); y) \wedge \mathrm{end}(y) = \mathrm{Sub}(v, \underset{\sim}{x}))\urcorner (\underset{\sim}{\eta}/v) \rightarrow \eta(x)),$$

where $\eta \in L_{\mathrm{PA}}$ is some formula less than b.

The above formula is obviously equivalent in $\mathrm{I}\Sigma_1$ to the following formula:

$$\theta_1 := \forall x\, \forall y\, \forall z\, (\ulcorner \mathrm{Prf}(\mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA})); y) \wedge \mathrm{end}(y) = \mathrm{Sub}(v, \underset{\sim}{x})\urcorner (\underset{\sim}{\eta}/v) \rightarrow \eta(x)),$$

where Q' is such that $\sup \alpha_{\underset{\sim}{z}}^{Q'} = \alpha$ and $(\varepsilon_\alpha)_z^P \Rightarrow_0 \varepsilon_{\alpha_{\underset{\sim}{z}}^{Q'}+1}$ for $\alpha \in \mathit{Lim}$ (cf. Note III.1.6(a)).

Since θ_1 is an expression built up, in a concrete number of steps, of η and $\underset{\sim}{\eta}$ and $\underset{\sim}{\alpha}$ smaller than b, it follows that, if c is large enough, say if $c \geq c_2 \geq c_1$ then $\theta_1 < a$ (cf. Note II.2.12). For the same reason we can assume that if $c \geq c_2$ then the proof of the equivalence $\theta \Leftrightarrow \theta_1$ has width less than a. Hence, by the soundness Lemma II.1.10, $A \models \theta$ iff $A \models \theta_1$.

We show that $A \models \theta_1$. Since $\theta_1 < a$, to proof $A \vdash \theta_1$ it is enough to show that for all $x, y, z < a_0$

$$A \models \ulcorner \mathrm{Prf}(\mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA})); y) \wedge \mathrm{end}(y) = \mathrm{Sub}(v, \underset{\sim}{x})\urcorner (\underset{\sim}{\eta}/v) \text{ implies } A \models \eta(x).$$

Since the formula to the right of the first sign $\models$ is of class Σ_1, it is upwards absolute (cf. Note II.2.11). Hence it suffices to prove that

$$\mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA})) \vdash_y \eta(\underset{\sim}{x}) \text{ implies } A \models \eta(x) \quad \text{for all } x, y, z < a_0.$$

Take $x, y, z < a_0$ and assume that $\mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA})) \vdash_y \eta(\underset{\sim}{x})$. In particular, $\mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA})) \vdash_{a_0} \eta(\underset{\sim}{x})$. By 2.1, there exists $B \subseteq A \setminus \{a_0\}$ which is an $\varepsilon_{\alpha_{\underset{\sim}{z}}^{Q'}} + 1$, P-large a_1-skeleton.

By the inductive assumption, if we put $b' = (a_1)^{1/c}$ we obtain $B \models_{b'} \mathrm{Ax}(\mathscr{R}_{\alpha_{\underset{\sim}{z}}^{Q'}}(\mathrm{PA}))$. Hence under the additional assumption that $(a_0)^c \leq a_1$ we obtain $B \models \eta(\underset{\sim}{x})$. Since A is diagonally indiscernible for η^* it follows that $A \models \eta(x)$ and the conclusion of the lemma follows. Observe that for $b = a^{1/c} < 2$ the conclusion is trivial.

Hence to finish the proof it is enough to choose a concrete constant c_3 such that the condition $a \geq c \geq c_3$ implies that $(a_0)^c \leq a_1$. (Then we put finally $c = \max(c_1, c_2, c_3)$.) By II.2.11 there exists a concrete constant $c_3 \geq 4$ such that every

a-skeleton A with $a \geqslant c_3$ is x^x-scattered apart a few final elements. Hence if $a \geqslant c \geq c_3$ then $a_0^c < a_1$ and the lemma is proved. □

1.3. Note. Lemma IV.1.2 suggests the following question: does the polynomial reduction of $\mathscr{R}_\alpha(\mathrm{PA})$ to $\mathrm{I}(\varepsilon_\alpha)$ exist, in symbols, does $\mathscr{R}_\alpha(\mathrm{PA}) \subseteq^{\mathrm{pol}} \mathrm{I}(\varepsilon_\varepsilon)$ hold (for an explanation of the notation see II.1.9).

In the case of an affirmative answer, Lemma III.1.5 immediately yields Lemma IV.1.2. Indeed the answer is positive, but to establish it we need Lemma IV.1.2. This can be established by an easy estimation of the length of the proof in $\mathrm{I}(\varepsilon_\alpha)$ in the proof of IV.1.1 below.

Proof of Theorem 1.1($\supseteq$). Let $n \geqslant 1$. For $T = \mathscr{R}_{<\alpha}(\mathrm{PA})$ Lemma II.2.6 reads as follows:

$$\mathrm{I}\Sigma_1 \vdash \mathrm{CR}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_n) \rightarrow \mathrm{R}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_n),$$

where the combinatorial reflection principle $\mathrm{CR}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_n)$ says:

$$\forall b\, \exists A\, [\text{“}A \text{ is a } \Sigma_{n-1}\text{-scattered } b\text{-skeleton”} \wedge A \vDash_b \mathscr{R}_{<\alpha}(\mathrm{PA})].$$

Hence it is enough to show that $\mathrm{I}(\varepsilon_\alpha) \vdash \mathrm{CR}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_n)$ for every natural n.

Now we fix n. From now on we work in the system $\mathrm{I}(\varepsilon_\alpha)$. For given b we choose a β less than α and a number b_1 such that

$$\forall \theta < b\, (\theta \in \mathscr{R}_{<\alpha}(\mathrm{PA}) \rightarrow \mathrm{Ax}(\mathscr{R}_\beta(\mathrm{PA})) \vdash_{b_1} \theta).$$

Obviously it is enough to construct a Σ_{n-1}-scattered b_1-skeleton A such that $A \vDash_{b_1} \mathrm{Ax}(\mathscr{R}_\beta(\mathrm{PA}))$. Assume that $P \in \Sigma_1$ defines in $\mathrm{I}\Sigma_1$ a $(\mathrm{B}^+)^2$-ε-system of sequences for $\varepsilon_{\beta+1}$.

Let c_0 be the number from 1.2. Put $a = (b_1)^{c_0}$ and $\gamma = \omega_a^{\varepsilon_\beta+1}$. Let $d = (H^{n-1})_\gamma^P(a)$. We define $B = \{(H^{n-1})^r(a)\colon (H^{n-1})^r(a) \leqslant d\}$. It is easy to check that B is $\omega_a^{\varepsilon_\beta+1}$, P-large and Σ_{n-1}-scattered. Hence by II.3.4(ii) there exists an $\varepsilon_\alpha + 1$, P-large a-skeleton $A \subseteq B$. Thus, by 1.2, $A \vDash_{b_1} \mathrm{Ax}\,(\mathscr{R}_\beta(\mathrm{PA}))$ and this finishes the proof. □

Since the proof of Theorem 1.1 is formalizable in $\mathrm{I}\Sigma_1$, III.1.12 shows the following corollary.

1.4. Corollary. *For all* $n \geqslant 1$, $\mathrm{I}\Sigma_1 \vdash \mathrm{CR}(\mathscr{R}_\alpha(\mathrm{PA}); \Sigma_n) \Leftrightarrow \mathrm{R}(\mathscr{R}_\alpha(\mathrm{PA}); \Sigma_n)$.

IV.2. The iterations of Ramsey's combinatorial principle

The aim of this section is to prove that the appropriate αth iterate of the Paris–Harrington principle, PH_Q^α, is independent of $\mathrm{I}(<\varepsilon_\alpha)$ or rather of $\mathscr{R}_{<\alpha}(\mathrm{PA})$ (by 1.1, $\mathrm{I}(<\varepsilon_\alpha) \equiv \mathscr{R}_{<\alpha}(\mathrm{PA})$).

Before we formulate exactly the principle PH_Q^α let us consider some ways of strengthening the principle PH.

Assume that $W(H, e, k, c)$ is some property of finite sets; H varies over finite sets, c is a fixed parameter. Then we define the following strengthening of Ramsey's combinatorial principle:

$$\mathrm{RP}[W] \equiv \forall e, k\, \exists r\, \exists M\, \forall F:[M]^e \to r\, \exists H \in \mathrm{Hom}(F)\, [|H| \geqslant \max(k, e+1) \wedge W(H, e, k, c)],$$

where $H \in \mathrm{Hom}(F)$ is a shorthand for $H \subseteq \mathrm{dom}(F) \wedge F\restriction[H]^e \equiv \mathrm{Const}$. Observe that

$$\mathrm{I}\Delta_0 + \exp \vdash \mathrm{RP}[\text{“}H \text{ is } \omega\text{-large”}] \equiv \mathrm{PH}.$$

Requiring H to be much larger we obtain a direct strengthening of PH. In the companion paper [24] it was proved that the principle RP[“H is $(\varepsilon_\alpha)_e$-large”] is independent of $\mathrm{I}(\varepsilon_\alpha)$ (for appropriate systems of notations). In fact, it is equivalent to $\mathrm{R}(\mathrm{I}(\varepsilon_\alpha); \Sigma_1)$ over $\mathrm{I}\Sigma_1$.

Now we fix a basic system Q of notations for λ in $\mathrm{I}\Sigma_1$ and of class Σ_1 and look for iterations of PH, formed in such a way that we require H to satisfy the combinatorial principle of some previous iterates. McAloon [18] defined in this way the combinatorial property $X \to^\alpha (m)^n_k$. It follows by recursion on Σ_1 formulas in PA that there exists a Σ_1 formula $X \to^\alpha (m)^n_k$ such that PA proves that for all $\alpha < \lambda$:

(i) $X \to^0 (m)^n_k \Leftrightarrow X \to_* (m)^n_k$ (Paris–Harrington property),

(ii) $X \to_\alpha (m)^n_k \Leftrightarrow \forall(F:[X]^n \to k)\, \exists H \in \mathrm{Hom}(F)$
$[(|H| \geqslant \max(n+1, m) \wedge \forall i \leqslant n\, (H \to^{\alpha_i} (1 + h_\omega)^{h_\omega}_{h_\omega})]$, for $\alpha > 0$,
where h_ω denotes the element h_{h_0} from H.

McAloon proved:

2.1. Theorem. *For every solution of* (i) *and* (ii),

$$\mathrm{PA} \vdash \forall x\, \exists y\, ([1, y] \to^\alpha (x+1)^x_x) \equiv \mathrm{R}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_1).$$

We consider here an iteration of PH having the simplest form.

2.2. Definition. $X \to_\alpha (\cdot)^n_k$ denotes an arbitrary Σ_1 formula that $\mathrm{I}\Sigma_1$ proves that for every $\alpha < \lambda$

(i) $X \to_0 (\cdot)^n_k \Leftrightarrow X \to_* (0)^n_k$,

(ii) $X \to_\alpha (\cdot)^n_k \Leftrightarrow \forall(F:[X]^n \to k)\, \forall i \leqslant n\, \exists H \in \mathrm{Hom}(F)$
$(H \to_{\alpha_i} (\cdot)^{h_1}_2)$, *for* $\alpha > 0$.

Finally, we define

$$\mathrm{PH}^\alpha_Q \equiv \forall n\, \forall k\, \exists M\, \forall(F:[M]^n \to 2)\, \forall i \leqslant n\, \exists H \in \mathrm{Hom}(F)(|H| \geqslant k \wedge H \to_{\alpha_i} (\cdot)^{h_1}_2).$$

2.3. Theorem. $\mathrm{I}\Sigma_1 \vdash \mathrm{PH}^\alpha_Q \Leftrightarrow \mathrm{R}(\mathscr{R}_{<\alpha}(\mathrm{PA}); \Sigma_l)$ *for* $\alpha > 0$.

The proof of Theorem 2.3($\leftarrow$) is a copy of the McAloon [18] purely syntactical proof of Theorem 2.1($\leftarrow$) and we omit it. The proof of Theorem 2.1($\rightarrow$) in [18] uses the construction of a theory which is a sort of iteration of the collection of all principles PH_n: n arbitrary. McAloon showed that this theory includes $\mathcal{R}_{<\alpha}(\mathrm{PA})$. The general scheme of the proof is very similar to the old proof of the Kreisel–Levy–Schmerl Theorem. The proof of Theorem 2.3($\rightarrow$) presented here is a sort of 'complete reduction in one step' as the proof in the previous section. This refined method applies as well to Theorem 2.1($\rightarrow$) but it is not known whether conversely.

Proof (of $\rightarrow$ in $\mathrm{I}\Sigma_1$). Assume PH^{α}_{Q}. By Lemma II.2.6 it is enough to show that $\mathrm{CR}(\mathcal{R}_{<\alpha}(\mathrm{PA}); \Sigma_1)$ which reads

$$\forall a\, \exists A\, (\text{“}A \text{ is an } a\text{-skeleton”} \wedge A \vDash_a \mathrm{Ax}(\mathcal{R}_{<\alpha}(\mathrm{PA}))).$$

It is difficult to do this directly for the hierarchy $\mathcal{R}_\beta(\mathrm{PA})$. We need a new hierarchy $\underline{\mathcal{R}}_\beta(\mathrm{PA})$ (or rather a new definition of the old hierarchy). We define $\mathrm{Ax}(\underline{\mathcal{R}}_{-1}(\mathrm{PA})) = \mathrm{Ax}(\mathrm{PA})$,

$$\begin{aligned}\mathrm{Ax}(\underline{\mathcal{R}}_\alpha(\mathrm{PA})) = \mathrm{Ax}(\mathrm{PA}) \cup \{\forall x\, \forall y\, \forall z\, (\ulcorner \mathrm{Prf}(\mathrm{Ax}\, \underline{\mathcal{R}}_{\alpha_z}; y) \wedge \mathrm{end}(y) \\ = \mathrm{Sub}(v, \underline{x})\urcorner (\underline{\eta}/v) \rightarrow \eta(x))\colon \eta \in L_{\mathrm{PA}}\} \quad \text{for } \alpha \geqslant 0.\end{aligned}$$

Finally, we put $\underline{\mathcal{R}}_\alpha(\mathrm{PA}) \equiv \mathrm{Cn}(\mathrm{Ax}(\underline{\mathcal{R}}_\alpha(\mathrm{PA})))$.

One can show that $\underline{\mathcal{R}}_\alpha(\mathrm{PA}) \equiv \mathcal{R}_\alpha(\mathrm{PA})$ by means of the Schmerl trick mentioned in the previous section.

We show that

$$\forall a\, \forall e\, \exists A\, [\text{“}A \text{ is an } a\text{-skeleton”} \wedge A \vDash_a \mathrm{Ax}(\underline{\mathcal{R}}_{\alpha_e}(\mathrm{PA}))\text{”}].$$

Take an arbitrary e and take an arbitrary $a \geqslant e$ greater than the constant c_0 from II.1.11 which guarantees that a-skeletons are $(x+1)^2$-scattered. Let b be an arbitrary number $\geqslant a + \max(2^a_4, c_1)$, where c_1 is greater than the constants from Lemmas II.4.2 and II.4.5.

By PH^{α}_{Q} there exists an M such that

$$\forall (F : [M]^a \rightarrow 2)\, \forall i \leqslant a\, \exists A \in \mathrm{Hom}(F)\, (|A| \geqslant b \wedge A \rightarrow_{\alpha_i} (\cdot)^{a_1}_{2}).$$

We put $F : [M]^a \rightarrow 2$ equal to the partition from II.4.5 which forces the existence of a-skeletons. We put $i = e$. Hence there exists an A such that $A \rightarrow_{\alpha_e} (\cdot)^{a_1}_{2}$ and $\lg^4[A \setminus \{a_0\}]$ minus its last a elements is an a-skeleton with $a_0 \geqslant c_1$.

Hence to complete the proof of the theorem it is enough to show the following claim for every $a < \lambda$.

Claim. *If $A \rightarrow_\alpha (\cdot)^{a_1}_{2}$, $a_0 \geqslant c_1$ and $A' = \lg^4[A \setminus \{a_0\}]$ minus its last a elements is an a-skeleton, then $A' \vDash_a \mathrm{Ax}(\underline{\mathcal{R}}_\alpha(\mathrm{PA}))$.*

To prove this claim we introduce more transparent notation. Let us denote $A \rightarrow_\alpha (\cdot)^{a_1}_{2}$ by $W(A, \alpha)$ and $\lg^4(x)$ by $G(x)$. The set B minus $\{b_0\}$ and minus its a

elements is denoted by B_a. We make the convention that if $\alpha = 0$ then $\alpha_x = -1$ for each x.

We define

(i) $$W(A, -1) \equiv |A| \geqslant a_0.$$

Now, by 2.2, $W(A, \alpha)$ implies that

$$\forall (F : [A]^{a_1} \to 2)\, \forall z \leqslant a_1\, \exists B \in \mathrm{Hom}(F)\, W(B, \alpha_z)].$$

By II.4.2 there exists a partition $P : [A \setminus \{a_0\}]^{a_1} \to 2$ such that if $B \in \mathrm{Hom}(P)$ then $\lg^4[B \setminus \{b_0\}]$ minus its last $\lg^4(a_1)$ elements is a $\lg^4(a_1)$-skeleton. Using the idea of the proof of II.4.5 we define $F : [A]^{a_1} \to 2$ having the same property.

Hence the following implication is valid:

(ii) $$\begin{aligned} &W(A, \alpha) \to \forall z \leqslant a_1\, \exists B \subseteq A \\ &\quad [W(B, \alpha_z^Q) \wedge \text{“}G[B_{G(a_1)}] \text{ is an } G(a_1)\text{-skeleton”}] \\ &\quad\quad \text{for all } \alpha,\ 0 \leqslant \alpha < \lambda. \end{aligned}$$

The claim will be proved as soon as we show the following lemma (we assume $W(A, \alpha)$, $G(x) = y \in \Sigma_1$ and c_0 is the constant from II.1.11).

2.4. Lemma ($\mathrm{I}\Sigma_1$). *If G is nondecreasing and $W(A, \alpha)$ has properties* (i) *and* (ii) *then the following implication is valid*:

$$W(A, \alpha) \wedge \text{“}G[A_a] \textit{ is an } a\textit{-skeleton”} \wedge a \geqslant c_0 \to G[A_a] \models_a \mathrm{Ax}(\mathscr{R}_\alpha(\mathrm{PA})).$$

Proof ($\mathrm{I}\Sigma_1$). Our proof is a refinement of the proof of Lemma 1.3. Fix an A and an α such that $W(A, \alpha)$. Let O be the set of ordinals of the form $(\cdots((\alpha_{a^1})_{a^2})\cdots)_{a^k}$, where $a^1, a^2, \ldots, a^k$ varies over all increasing sequences of elements less than $\max A$ (including the empty sequence, hence $\alpha \in O$). Observe that O is finite.

We show by induction on $\beta \in O$ that the following implication is valid:

$$\begin{aligned} &W(B, \beta) \wedge \text{“}G[B_a] \text{ is an } a\text{-skeleton”} \wedge a \geqslant c_0 \to \\ &\quad G[B_a] \models_a \mathrm{Ax}(\mathscr{R}_\beta(\mathrm{PA})) \quad \text{for all } B \subseteq A \text{ and all } a \leqslant \max A. \end{aligned}$$

This is obviously true for $\beta = -1$. To show the inductive step from all $\gamma < \beta$ which belong to O to $\beta \in O$, where $\beta \geqslant 0$, assume that $B \subseteq A$, $W(B, \beta)$ and $G[B_a]$ is an a-skeleton, where $c_0 \leqslant a < \max A$.

Let

$$\theta := \forall x\, \forall y\, \forall z\, (\ulcorner \mathrm{Prf}(\mathrm{Ax}\, \mathscr{R}_{\beta_z}; y) \wedge \mathrm{end}(y) = \mathrm{Sub}(v, \underset{\sim}{x})\urcorner (\eta/v) \to \eta(x))$$

be less than a. It is enough to show that $G[B_a] \models \theta$.

Since $G[B_a]$ is an a-skeleton it is enough to show that for all $x, y, z < \min G[B_a]$, $\mathrm{Ax}(\mathscr{R}_{\beta_z}(\mathrm{PA})) \vdash_y \eta(\underset{\sim}{x})$ implies that $G[B_a] \models \eta(\underset{\sim}{x})$ (cf. the proof of Lemma 1.3).

We know that $\min G[B_a]=G(b_1)$. Hence take $x, z<G(b_0)$ and assume that $\mathrm{Ax}(\mathscr{R}_{\beta_z}(\mathrm{PA}))\vdash_{G(b_1)}\eta(\underset{\sim}{x})$.

By (ii) it follows that there exists $C\subseteq B$ such that $W(C,\beta_z)$ and $G[C_{G(b_1)}]$ is an $G(b_1)$-skeleton. Hence by the inductive assumption $G[C_{G(b_1)}]\vdash_{G(b_1)}\mathrm{Ax}(\mathscr{R}_{\beta_z}(\mathrm{PA}))$. Thus by the Soundness Lemma $G[C_{G(b_1)}]\models\eta(x)$ and finally also $G[B_a]\models\eta(\underset{\sim}{x})$, because $\underset{\sim}{x}<G(b_0)$ and $\eta<a$. □

V. Transfinite induction in the language with a satisfaction class

In this chapter we extend the results of Chapter III beyond subsystems of the true first-order arithmetic.

We deal here with satisfaction classes. Given $M\models\mathrm{PA}$, by a full satisfaction class for M we mean a subset $S\subseteq M\times M$ of pairs $(\theta,\bar{a})$ of the form ⟨formula, sequence⟩ in the sense of M for which the usual Tarski condition on truth is satisfied (see Krajewski [15] for a more precise definition).

Let Sat(S)) be a Π_2 formula of the language $L_{\mathrm{PA}}(S)=L_{\mathrm{PA}}\cup\{S\}$ which describes these conditions. We set by convention $\varepsilon_{-1}=\omega$.

Let $\mathrm{I}(\varepsilon_\alpha;S)$ denote $\mathrm{I}\Sigma_1$ plus Sat(S) enriched with the scheme of transfinite induction up to ε_α for all formulas of $L_{\mathrm{PA}}(S)$. The theory $\mathrm{I}(\varepsilon_{-1};S)$ is usually denoted by PA(S).

All results of Chapter III can be extended to the case of $\mathrm{I}(\varepsilon_\alpha;S)$. For illustration we only show the following theorem.

1. Theorem. *Assume that P of class Σ_1 defines in $\mathrm{I}\Sigma_1$ a $(B^+)^2$-ε-system for $\varepsilon_{\varepsilon_{\alpha+1}}$. Then*

$$\mathrm{I}(\varepsilon_\alpha;S)\restriction\Pi_{n+2}\equiv P\text{-}\mathrm{TH}^n(<\varepsilon_{\varepsilon_{\alpha+1}})\quad\textit{for all } n.$$

Corollary. $\mathrm{I}(\varepsilon_\alpha;S)\restriction L_{\mathrm{PA}}\equiv\mathrm{I}(<\varepsilon_{\varepsilon_{\alpha+1}})$.

The simplest case of corollary for $\alpha=-1$, i.e. the statement $\mathrm{PA}(S)\restriction L_{\mathrm{PA}}\equiv\mathrm{I}(<\varepsilon_{\varepsilon_0})$, was proved in [11], in a direct manner, without using the logical properties of a-skeleton. The method used there relies on a combinatorial construction of a full satisfaction class S over recursively saturated models for $\mathrm{I}(<\varepsilon_{\varepsilon_0})$.

Now to prove Theorem 1, we construct in $P\text{-}\mathrm{TH}^n(<\varepsilon_{\varepsilon_{\alpha+1}})$ some fragments of satisfaction classes, so-called ε_α-large a, S-skeletons and we will use their logical properties.

Obviously the essential difficulty lies in showing the inclusion $\subseteq$ in Theorem 1. The opposite inclusion follows from the following observations: first $\mathrm{I}(\varepsilon_\alpha;S)\equiv\mathrm{I}(<\varepsilon_{\alpha+1};S)$. Moreover, we easily show by induction on γ that

$$\mathrm{I}(<\varepsilon_{\alpha+1};S)\vdash\forall\gamma\,[-1\leq\gamma<\omega_n^{\varepsilon_\alpha+1}\rightarrow\forall\theta\,S(\mathrm{Ind}(\theta,\varepsilon_\gamma))]\quad\text{for all } n.$$

Hence $\mathrm{I}(<\varepsilon_{\varepsilon_{\alpha+1}})\subseteq\mathrm{I}(\varepsilon_\alpha;S)$ and Theorem 1($\supseteq$) follows.

Proof of Theorem 1($\subseteq$). We first assume that

$$\mathrm{I}(\varepsilon_\alpha; S) \vdash_m \forall x\, \exists y\, \varphi(x, y),$$

where φ is of class Π_{n+2}. By Σ_1-completeness

$$\mathrm{I}\Sigma_1 \vdash (\mathrm{I}(\varepsilon_\alpha; S) \vdash_{\underline{m}} \ulcorner \forall x\, \exists y\, \varphi(x, y) \urcorner).$$

Then we want to work in P-$\mathrm{TH}^n(<\varepsilon_{\varepsilon_{\alpha+1}})$. By Corollary III.1.9 we have freedom in choosing P, because P-$\mathrm{TH}^n(<\varepsilon_{\varepsilon_{\alpha+1}})$ is the same theory for all $(\mathrm{B}^+)^2$-ε-systems P in $\mathrm{I}\Sigma_1$.

For our purposes we assume that P is a $(\mathrm{B}^+)^2$-ε-system in $\mathrm{I}\Sigma_1$ such that there exists a $Q \in \Sigma_1$ such that

(1) Q is a $(\mathrm{B}^+)^2$-ε-system (in $\mathrm{I}\Sigma_1$) and

(2) $(\varepsilon_\alpha)^P = \varepsilon_{\alpha^Q_{\underline{n}}+1}$ for all $\alpha \in Lim \setminus \{0\}$ (in $\mathrm{I}\Sigma_1$).

Having chosen such a P (it exists by I.2.2) we work in P-$\mathrm{TH}^n(<\varepsilon_{\varepsilon_{\alpha+1}})$.

Take an $a \geqslant m^c$, for c to be chosen later. Our main aim is the same as in the proof of III.1.4 (where we worked in P-$\mathrm{TH}^n(<\varepsilon_{\alpha+1})$). Namely to construct a Σ_n-scattered $\underline{m}^c$-skeleton C with $\min C \geqslant a$ such that $C \vDash \ulcorner \forall x\, \exists y\, \varphi(x, y) \urcorner$, which implies that $\forall x \leqslant a\, \exists y \leqslant c_1\, \varphi(x, y)$ and finishes the proof.

Let $\beta(m, c) = \varepsilon_\gamma$, where $\gamma = \omega^{\varepsilon_\alpha+1}_{\underline{m}^c}$. First we construct in an obvious way a Σ_n-scattered $\beta(m, c)$-large set A with $\min A = a$.

Now we roughly describe the remaining three steps. The second step is to choose some special $\omega^{\varepsilon_\alpha+1}_{\underline{m}^c}$, Q-large set $B \subseteq A$ called an S-skeleton.

In the third step we construct an $\varepsilon_\alpha + 1$, Q-large set $C \subseteq B$ which in a relativized manner is an '$\underline{m}^c$, S-skeleton', i.e., it is an S-skeleton and we have well defined $C \vDash \eta$ for $\eta \in L_{\mathrm{PA}}(S)$, $\eta < \underline{m}^c$.

Moreover, for such S-skeletons we show that $C \vDash \mathrm{Sat}(S)$.

In the last (fourth) step we show how Lemma III.1.5 should be modified to obtain the relativized version of III.1.5 saying roughly that $C \vDash_{\underline{m}^c} \mathrm{Ax}(\mathrm{I}(\varepsilon_\alpha; S))$ for large enough c.

To realize the second step we need some definitions.

2. Definition. A is $\omega \dotplus \beta$, *Q-large* iff there exist sets B^1 and B^2 such that

(1) $\max B^1 = \min B^2$ and $B = B^1 \cup B^2$,

(2) B^1 is β-large and B^2 is ω-large.

If $\omega \dotplus \beta$ is well defined then obviously B is $\omega \dotplus \beta$, Q-large in the above sense iff B is $\omega \dotplus \beta$, Q-large in the usual sense.

Let B be a set. We define $b^\gamma = \max_b\{B \cap [b, \max B]$ is γ-large$\}$ and $B^\gamma = \{b \in B : b \leqslant b_\gamma\}$.

3. Definition. We say that B is an *S-skeleton* iff b is ω-large and for every $a \in B^\omega$ the set $B \cap (a, \max B]$ is an a-skeleton.

If B is an S-skeleton then we define

$$B \vDash S(\theta, \bar{x}) \Leftrightarrow \theta < b^{\omega} \wedge \bar{x} < b^{\omega} \wedge B \vDash \theta(\bar{x}).$$

Let us set $\lambda = \varepsilon_{\varepsilon_{\alpha+1}}$.

4. Lemma. *There exists a concrete constant c such that for all β satisfying $\beta < \lambda$ and for each ε_β, P-large set A with $a_0 \geq c$ there exists an $\omega \dot{+} \beta$, Q-large S-skeleton $B \subseteq A$.*

This lemma completes the second step. Namely, if we put $\beta = \beta(m, c)$ in 4, then we obtain the desired $\omega \dot{+} \omega_{m^c}^{\varepsilon_\alpha+1}$-large S-skeleton $B \subseteq A$.

Proof. Take for c ($c \geq 3$) the constant from II.2.13 for $F(x) = x + 1$. Thanks to Lemma I.2.6 on finitization we can freely argue by induction on α.

If $\beta = 0$ then by II.3.4(i) there exists an a_0-skeleton $B' \subseteq A$. Since $b'_0 > a_0$ the set $B = \{a_0\} \cup \{b'_0, \ldots, b'_{a_0+1}\}$ is the required S-skeleton.

Now we show the inductive step from all appropriate $\gamma < \beta$ to β for $\beta \geq 0$.

Let A be an ε_β, P-large with $a_0 \geq c$. By Lemma III.2.1 there exists an $\varepsilon_{\alpha^Q_{a_0}} + 1$, P-large a_0-skeleton $A' \subseteq A$. By the choice of c, $A' \subseteq A \setminus \{a_0\}$.

By the inductive assumption there exists an $\omega \dot{+} \alpha^Q_{a_0}$, Q-large S-skeleton $B' \subseteq A'$. We put $B = \{a_0\} \cup B'$. Hence B is $\omega \dot{+} \alpha^Q_{a_0} + 1$, Q-large, i.e., B is $\omega + \alpha$, Q-large if α is nonlimit. If α is limit we infer first that B is $\omega \dot{+} \alpha^Q_{a_0}$, Q-large. Hence also B is $\omega \dot{+} \alpha$, Q-large. Moreover, $B \cap (a_0, \max B] = B' = A'$ is an a_0-skeleton. Hence B is also an S-skeleton. $\square$

To realize the third step we need a few auxiliary definitions. Assume that $\eta \in L_{\mathrm{PA}}(S)$ and A is an S-skeleton. Let η^{*A} define a Δ_0-formula obtained from η^* by substitutions of all appearances of $S(\theta, \bar{x})$ by $A \vDash \theta(\bar{x})$.

5. Definition. An S-skeleton A is called an *a, S-skeleton* iff A is $\omega \dot{+} a$-large, $a \leq a_0$, and A^{ω} is a set of diagonally indiscernible elements for all relations $\mathrm{Tr}_0(\eta^{*A}(\bar{x}, \bar{y}))$, where η ranges over all generalizations of all formulas $\eta \in L_{\mathrm{PA}}(S)$ less than a. Moreover, we assume that A^{ω} is a set of indiscernible elements for $\mathrm{Tr}_0(\eta^{*A}(\bar{x}, \bar{y}))$ for all those η as above which are sentences.

6. Definition. If A is an a, S-skeletons and $\eta \in L_{\mathrm{PA}}(S)$, $\eta < a$, then we define

$$A \vDash \eta(\bar{x}) \Leftrightarrow \mathrm{Tr}_0(\eta^{*A}(\bar{x}, a^{\omega+n-1}, \ldots, a^{\omega})), \quad \text{where } n = \text{a.r.}(\eta).$$

Since $A \setminus \{a_0\}$ is an a_0-skeleton the above definition agrees with the usual definition of the relation $A \vDash \eta(\bar{x})$ for $\eta \in L_{\mathrm{PA}^-}$. Now, we relativize without difficulty all lemmas of Chapter II. The relativized versions are marked, by the letter S. We now describe successively these versions. Passing to II.1.9(i)–(ii)^S, II.1.10^S, II.1.12^S consists in the following changes in the corresponding lemmas and proofs: the set A is replaced by an S-skeleton A, the a-skeleton A is replaced by an a, S-skeleton A.

The relativized version of Lemma II.1.9(iii) has the following form.

7. Lemma ($\mathrm{I}\Delta_0+\exp$). *If A is an a, S-skeleton, $A^\omega=\{a_0,\ldots,a_k\}$, $\theta\in L_{\mathrm{PA}}(S)$, $\theta(\bar{x},\bar{y})<a$, then*

$$\forall\bar{x}<a^{\omega+r+1}[A\vDash\forall\bar{y}\,\theta(\bar{x},\bar{y})\Leftrightarrow\forall\bar{y}<a^{\omega+r}A\vdash\theta(\bar{x},\bar{y})],$$

where r is a number such that either $\mathrm{a.r.}(\theta)\leqslant r\leqslant k-1$, *or $r=k$ and $\bar{x}$ is empty.*

(Proof without essential changes.)

The relativized version of Lemma II.1.12 has the following form.

8. Lemma ($\mathrm{I}\Delta_0+\exp$). *There exists a constant c_0 such that for every a, S-skeleton A with $a\geqslant c_0$, $A\vDash_a\mathrm{Ax}(\mathrm{PA}(S))$.*

Proof. Obviously, it is enough to show that there exists c_0 such that for every a, S-skeleton A with $a\geqslant c_0$, $A\vDash\mathrm{Sat}(S)$.

Assume temporarily that $c_0>\ulcorner\mathrm{Sat}(S)\urcorner$; and appropriate value of c_0 will be chosen in the course of the proof. Assume now that A is an a, S-skeleton with $a\geqslant c_0$.

We only show that if c_0 is large enough then

$$A\vDash\ulcorner\forall\theta\in L_{\mathrm{PA}}(S)\,\forall\bar{b}\in\mathrm{Seq}\,\forall v\in\mathrm{Zm}\,(S(\forall v\,\theta,\bar{b})\Leftrightarrow\forall x\,S(\theta,\bar{b}(x/v)))\urcorner.$$

By 7. it is enough to show that

$$\forall\theta,\bar{b},v<a_0\,(A\vDash\theta\in L_{\mathrm{PA}}(S)\wedge\bar{b}\in\mathrm{Val}\wedge v\in\mathrm{Zm}\rightarrow$$
$$S(\forall v\,\theta,\bar{b})\Leftrightarrow\forall x\,S(\theta,\bar{b}(x/v)).$$

The formulas: $\theta\in L_{\mathrm{PA}}(S)$, $\bar{b}\in\mathrm{Val}$, $v\in\mathrm{Zm}$ are of class $\Delta_0(2^x)$ and hence by II.2.14^S, when c_0 is large enough, these formulas are A^ω absolute for $\theta,\bar{b},v<a_0$. We now fix $\theta,\bar{b},v<a_0$ such that $\theta\in L_{\mathrm{PA}}(S)$, $\bar{b}\in\mathrm{Val}$, $v\in\mathrm{Zm}$. Hence it is enough to show

(1) $\quad A\vDash S(\forall v\,\theta,\bar{b})\Leftrightarrow A\vDash\forall x\,S(\theta,\bar{b}(x/v))$.

By the definition of S-skeletons

$$A\vDash S(\forall v\,\theta,\bar{b})\Leftrightarrow A\vDash\forall v\,\theta(\bar{b})\Leftrightarrow A\backslash\{a_0\}\vDash\forall v\,\theta(\bar{b}).$$

Since A is an S-skeleton and hence $A\backslash\{a_0\}$ is an a_0-skeleton it follows by II.1.9 that

$$A\vDash\forall v\,\theta(\bar{b})\Leftrightarrow\forall x<a_1\,A\vDash\theta(\bar{b}(x/v)).$$

Since A is $\omega\dotplus a$-large it follows that $a_1\leqslant a^\omega$ and hence

$$\forall x<a_1\,A\vDash\theta(\bar{b}(x/v))\Leftrightarrow\forall x<a_1\,A\vDash S(\theta,\bar{b}(x/v)).$$

Therefore

$$A \models S(\forall v\theta, \bar{b}) \Leftrightarrow \forall x < a_1\, A \models S(\theta, \bar{b}(x/v))$$

and by Lemma 7 the proof is finished. $\square$

To end the third step we consider the relativized version of II.3.4(ii) which has the following form.

9. Lemma. *If A is an $\omega \dot{+} \omega_a^{\varepsilon_\alpha+1}$-large S-skeleton, $a_0 \geq a \geq 3$, then there exists an $\omega \dot{+} \varepsilon_\alpha + 1$-large a, S-skeleton $B \subseteq A$.*

Since A^ω is $\omega_a^{\varepsilon_\alpha+1}$-large we can repeat the whole reasoning without essential changes and we obtain a $B' \subseteq A^\omega$ such that B' is $\varepsilon_\alpha + 1$-large and is a set of diagonally indiscernible elements over $\mathrm{Tr}(\eta^{*A}(\bar{x}, \bar{y}))$ for all $\eta < a$.

Let $B = B' \cup (A \setminus A^\omega)$. Then $B^\omega = B'$ is a set of diagonally indiscernible elements over $\mathrm{Tr}(\eta^{*B}(\bar{x}, \bar{y}))$ and is suitably indiscernible. Moreover, B is obviously an S-skeleton and hence B is an a, S-skeleton. $\square$

Finally, taking the $\omega \dot{+} \omega_{\underline{m}^c}^{\varepsilon_\alpha+1}$-large S-skeleton $B \subseteq A$ from the end of the second step and applying Lemma 9 we obtain an $\omega \dot{+} \varepsilon_\alpha + 1$-large $\underline{m}^c$, S-skeleton $C \subseteq A$. Obviously C is Σ_n-scattered and as we know we have to show that $C \models \ulcorner \forall x\, \exists y\, \varphi(x, y) \urcorner$.

We pass to the fourth step. Since $\mathrm{I}(\varepsilon_\alpha; S) \vdash_{\underline{m}^c} \ulcorner \forall x\, \exists y\, \varphi(x, y) \urcorner$ and the relativized version of the Soundness Lemma is valid, to finish the proof of Theorem 1 it is enough to show that $C \models_{\underline{m}^c} \mathrm{Ax}(\mathrm{I}(\varepsilon_\alpha; S))$ for large enough c. Hence it is enough to show the relativized version of III.1.5.

Since all lemmas which were needed to prove III.1.5 can be easily relativized, we only state the relativized version, omitting the description of changes in the proof.

10. Lemma. *There exists a concrete number c such that for all α satisfying $\alpha = -1 \vee \alpha < \lambda$ and for each $\omega \dot{+} \varepsilon_\alpha + 1$, P-large a, S-skeleton A,*

$$A \models_{a^{1/c}} \mathrm{Ax}(\mathrm{I}(\varepsilon_\alpha; S)).$$

References

[1] H. Bachmann, Transfinite Zahlen (Springer, Berlin, 1959).

[2] W. Buchholz and S. Wainer, Provably computable functions and the fast growing hierarchy, Contemporary Math. 65, (1987) 179–198.

[3] S. Buss, Bounded Arithmetic, (Bibliopolis, Napoli, 1986).

[4] S. Feferman, Transfinite recursive progressions of axiomatic theories, J. Symbolic Logic 27 (1962) 259–316.

[5] S. Feferman, Systems of predicative analysis, J. Symbolic Logic 33 (1968) 193–220.

[6] S. Feferman, Hilbert's program relativized: proof-theoretical and foundational reductions, J. Symbolic Logic 53 (1988) 364–384.

[7] G. Gentzen, Beweisbarkeit und unbeweisbarkeit von Aufangsfählen der transfiniten Induktion in der reinen Zahlentheorie, Math. Ann. 119 (1943) 140–161.

[8] A. Grzegorczyk, Some classes of recursive functions, Rozprawy matematyczne, no. 4, Warsaw, (1953).
[9] P. Hajek and J. Paris, Combinatorial principles concerning approximations of functions, Arch. Math. Log. 26 (1987) 13–28.
[10] H. Kotlarski and Z. Ratajczyk, Inductive full satisfaction classes, Ann. Pure Appl. Logic 47 (1990) 199–223.
[11] H. Kotlarski and Z. Ratajczyk, More on induction in the language with a satisfaction class. Z. Math. Logik, Grundlag Math. 36 (1990) 441–453.
[12] J. Ketonen and R. Solovay, Rapidly growing Ramsey functions, Ann. of Math. 11 (1981) 267–314.
[13] G. Kreisel, On the interpretation of nonfinitist proofs II, J Symbolic Logic 17 (1952) 43–58.
[14] G. Kreisel and A. Levy, Reflection principles and their use for establishing the complexity of axiomatic systems, Z. Math. Logik Grundlag. Math. 14 (1968) 97–142.
[15] S. Krajewski, Non-standard satisfaction classes, in: Set Theory and Hierarchy Theory, Lecture Notes in Math. 537 (Springer, Berlin, 1976) 121–145.
[16] R. Kurata, Paris–Harrington theory and reflection principles, Saitama Math. J. 2 (1984) 33–45.
[17] M.H. Löb and S.S. Wainer, Hierarchies of number-theoretic functions, Arch. Math. Logik 13 (1970) 39–51, 97–113.
[18] K. McAloon, Paris–Harrington incompleteness and progressions of theories, Proc. Symp. Pure Math. 42 (1985) 447–460.
[19] J. Mycielski, Locally finite theories, J. Symbolic Logic 51 (1986) 59–62.
[20] J. Paris, A hierarchy of cuts in models of arithmetic, in: Model Theory of Algebra and Arithmetic, Lecture Notes in Math. (Springer, Berlin, 1980) 312–337
[21] J. Paris and L. Harrington, Incompleteness in PA, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 1133–1142.
[22] Z. Ratajczyk, Satisfaction classes and combinatorial sentences independent from PA, Z. Math. Logik Grundlag. Math. 28 (1982) 149–165.
[23] Z. Ratajczyk, A combinatorial analysis of functions provably recursive in $I\Sigma_n$, Fund. Math. 130 (1988) 191–213.
[24] Z. Ratajczyk, Arithmetical transfinite induction and hierarchies of functions, Fund. Math. 141, (1992) 1-20.
[25] Z. Ratajczyk, Some remarks on Paris–Harrington theorem, in preparation.
[26] H.E. Rose, Functions and hierarchies, Oxford Logic Guides 9 (1984) 189.
[27] U. Schmerl, A fine structure generated by reflection formulas over primitive recursive arithmetic, in: M. Boffa, K. McAloon and D. Van Dalen, eds., Logic Colloquium 78 (North-Holland, Amsterdam, 1980) 335–350.
[28] D. Schmidt, Built-up systems of fundamental sequences and hierarchies of number-theoretic functions, Arch. Math. Logik 18 (1976) 47–53.
[29] D. Schmidt, Postscript to 'Built-up systems . . . ', Arch. Math. Logik 18 (1977) 145–146.
[30] C. Smorynski, The incompleteness theorems, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 821–865.
[31] G. Takeuti, Proof Theory (North-Holland, Amsterdam, 1987).
[32] S. Wainer, Ordinal recursion, and a refinement of the extended Grzegorczyk hierarchy, J. Symbolic Logic 37 (1972) 281–292.